



HAL
open science

Strong regularization by Brownian noise propagating through a weak Hörmander structure

Paul-Eric Chaudru de Raynal, Igor Honoré, Stephane Menozzi

► **To cite this version:**

Paul-Eric Chaudru de Raynal, Igor Honoré, Stephane Menozzi. Strong regularization by Brownian noise propagating through a weak Hörmander structure. 2018. hal-01906576v1

HAL Id: hal-01906576

<https://hal.science/hal-01906576v1>

Preprint submitted on 26 Oct 2018 (v1), last revised 25 Sep 2020 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Strong regularization by Brownian noise propagating through a weak Hörmander structure

Paul-Éric Chaudru de Raynal*, Igor Honoré† and Stéphane Menozzi‡

October 26, 2018

Abstract

We establish strong uniqueness for a class of degenerate SDEs of weak Hörmander type under suitable Hölder regularity conditions for the associated drift term. Our approach relies on the Zvonkin transform which requires to exhibit good smoothing properties of the underlying parabolic PDE with *rough*, here Hölder, drift coefficients and source term. Such regularizing effects are established through a perturbation technique (forward parametrix approach) which also heavily relies on appropriate duality properties on Besov spaces.

For the method employed, we exhibit some sharp thresholds on the Hölder exponents for the strong uniqueness to hold.

1 Introduction

1.1 Statement of the problem and related results

In this work, we aim at establishing a strong well posedness result outside the classical Cauchy-Lipschitz framework for the following degenerate Stochastic Differential Equation (SDE) of Kolmogorov type:

$$\begin{aligned} d\mathbf{X}_t^1 &= \mathbf{F}_1(t, \mathbf{X}_t^1, \dots, \mathbf{X}_t^n)dt + \sigma(t, \mathbf{X}_t^1, \dots, \mathbf{X}_t^n)dW_t, \\ d\mathbf{X}_t^2 &= \mathbf{F}_2(t, \mathbf{X}_t^1, \dots, \mathbf{X}_t^n)dt, \\ d\mathbf{X}_t^3 &= \mathbf{F}_3(t, \mathbf{X}_t^2, \dots, \mathbf{X}_t^n)dt, \\ &\vdots \\ d\mathbf{X}_t^n &= \mathbf{F}_n(t, \mathbf{X}_t^{n-1}, \mathbf{X}_t^n)dt, \end{aligned} \quad t \geq 0, \quad (1.1)$$

where, $(W_t)_{t \geq 0}$ stands for a d -dimensional Brownian motion on some filtered probability space $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ and for all $i \in \llbracket 1, n \rrbracket^1$, $t \geq 0$ the component \mathbf{X}_t^i is \mathbb{R}^d -valued as well (i.e. $\mathbf{X}_t \in \mathbb{R}^{nd}$). We suppose that the $(\mathbf{F}_i)_{i \in \llbracket 2, n \rrbracket}$ satisfy a kind of weak Hörmander condition, i.e. the matrices $(D_{\mathbf{x}_{i-1}} \mathbf{F}_i(t, \cdot))_{i \in \llbracket 2, n \rrbracket}$ have full rank. However, the coefficients $(\mathbf{F}_i)_{i \in \llbracket 2, n \rrbracket}$ can be rather *rough* in their other entries, namely, Hölder continuous. We assume as well that the diffusion coefficient σ is bounded from above and below and spatially Lipschitz continuous.

For a system of Ordinary Differential Equation (ODE) it may be a real challenge to prove the well posedness outside the Lipschitz framework (see *e.g.* [DPL89]) and, as shown by Peano's example, uniqueness may fail as soon as the drift of the system of interest is only Hölder continuous. For an SDE, the story is rather different since the presence of the noise may allow to restore well posedness. Such a phenomenon, called *regularization by noise* (see the Saint Flour Lecture notes of Flandoli [Fla11] and the references therein for an overview of the topic), has been widely studied since the pioneering unidimensional work of Zvonkin [Zvo74] and its generalization to the multi-dimensional setting by Veretennikov [Ver80] where SDEs driven by a *non-degenerate*

*Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LAMA, 73000 Chambéry, France. pe.deraynal@univ-smb.fr

†Laboratoire de Modélisation Mathématique d'Evry (LaMME), Université d'Evry Val d'Essonne, 23 Boulevard de France 91037 Evry. igor.honore@univ-evry.fr

‡Laboratoire de Modélisation Mathématique d'Evry (LaMME), Université d'Evry Val d'Essonne, 23 Boulevard de France 91037 Evry, France and Laboratory of Stochastic Analysis, HSE, Shabolovka 31, Moscow, Russian Federation. stephane.menzio@univ-evry.fr

¹We will use throughout the paper the notation $\llbracket \cdot, \cdot \rrbracket$ for integer intervals.

Brownian noise and a bounded drift are shown to be well-posed. We mean by *non-degenerate* that the noise has the same dimension as the underlying system on which it acts.

Let us mention, among others, and still within the *non-degenerate* setting, the works of Krylov and Röckner [KR05] ($L_q - L_p$ drift), Zhang [Zha10] ($L_q - L_p$ drift and weakly Lipschitz diffusion matrix) and also Fedrizzi and Flandoli [FF11] ($L_q - L_p$ and Hölder drift).

The crucial assumption, shared by all the aforementioned results, is the non-degeneracy condition assumed on the noise added in the considered system. A possible approach to relax this hypothesis was proposed by Veretennikov in [Ver83], where the author extended the result in [Ver80] to some specific case of the considered chain (1.1) for $n = 2$. In comparison with our setting, the author does not impose any non-degeneracy condition on $D_{\mathbf{x}_1} \mathbf{F}_2(t, \cdot)$. The to price pay is anyhow that all coefficients (i.e. with the notations of (1.1) $\mathbf{F}_1, \mathbf{F}_2, \sigma$) need to be twice continuously differentiable functions with bounded derivatives w.r.t the degenerate component, meaning that no regularization by noise is investigated in the degenerate direction. More generally, it is useless to expect a generalization of the previous results without any additional assumption: we can benefit from the regularization by noise phenomenon only in the directions submitted to the noise.

In our current framework, the non-degeneracy assumption on the Jacobian $(D_{\mathbf{x}_{i-1}} \mathbf{F}_i(t, \cdot))$, $i \in \llbracket 2, n \rrbracket$ precisely allows the noise to propagate through the chain passing from the i^{th} to the $(i+1)^{\text{th}}$ level thanks to the drift, hence leading to a propagation of the noise in the whole considered space. The main idea is then to take advantage of this particular propagation, known as *weak Hörmander setting* (in reference to the work of Hörmander on hypoelliptic differential operator [Hör67]), to restore strong well-posedness under our current Hölder framework. This feature has already been considered in the literature for the system (1.1) in the particular case $n = 2$, see the works of Chaudru de Raynal [CdR17], Wang and Zhang [WZ16], Fedrizzi, Flandoli, Priola and Vovelle [FFPV17], Zhang [Zha18]. In any cases, in addition to the weak Hörmander structure, the regularity of the drift with respect to (w.r.t.) the second space (and hence degenerate) argument is required to be of regularity index superior or equal (depending on the work) to $2/3$ (critical Hölder index or critical weak differentiation index). As a generalization of these results, we prove in this paper that strong well-posedness holds as soon as each drift component \mathbf{F}_i is β_j -Hölder continuous in the j^{th} variable for some $\beta_j \in \left((2j-2)/(2j-1), 1 \right]$ so that we recover the critical index mentioned above when $j = 2$.

One may hence wonder why our thresholds are greater than the ones obtained in the non degenerate framework and if they are sharp. These questions are, in fact, a bit involved in our current framework of *strong* regularization by noise. In order to understand what is expectable or not, let us first consider the case of weak regularization by noise (i.e. weak well-posedness of ODE perturbed by noise). Thanks to Delarue and Flandoli's result [DF14], Chaudru de Raynal [CdR18] and then Chaudru de Raynal and Menozzi [CdRM17] have proposed a class of counter-examples to weak well posedness for the chain (1.1) when $n = 2$ and in full generality, respectively. This class relies the Peano's example and, in its simplest form, says the following: weak uniqueness fails for the equation,

$$\dot{Y}_t = \text{sgn}(Y_t)|Y_t|^\alpha dt + d\mathcal{W}_t, \quad Y_0 = 0, \quad (1.2)$$

where \mathcal{W} is a continuous symmetric noise with self-similarity index $\gamma > 0$ as soon as $\alpha < 1 - 1/\gamma$. The point in such counter-examples, consists in comparing the maximal solutions of the ODE with dynamics $\dot{y}_t = \text{sgn}(y_t)|y_t|^\alpha$ starting from 0 (which is precisely the singular point of the ODE), which writes as $\pm ct^{1/(1-\alpha)}$, and the typical magnitude of the noise added in the system, which is t^γ . For the noise to dominate in small time, in order to leave the point where uniqueness fails for the deterministic equation, one must take $\gamma < 1/(1-\alpha) \Leftrightarrow \alpha > 1 - 1/\gamma$. Moreover, concerning our setting, building ad hoc version of the system (1.2) including possible iterated (in time) integral of order l of the solution (so that the maximal sol=write $ct^{(l\alpha+1)/(1-\alpha)}$), Chaudru de Raynal and Menozzi show that weak uniqueness may fail as soon as the condition $\alpha > (\gamma - 1)/(l + \gamma)$ is not satisfied (see Lemma 4.1 in [CdR18] and Proposition 4 in [CdRM17]).

It is hence clear that in our current degenerate framework the increasing value of the critical Hölder index can be understood as the price to pay to balance the degeneracy of the noise. Indeed, the way the noise is allowed to propagate here implies a loss of the fluctuations at each level of the chain and therefore, a loss of its regularization property.

A good manner to understand this phenomenon is the following case. Given a level i in $\llbracket 2, n \rrbracket$ of the chain (1.1) and a variable \mathbf{x}_j , j in $\llbracket i, n \rrbracket$, assume that $\mathbf{F}_1 \equiv 0$, $\sigma \equiv \text{Id}$, $\mathbf{F}_k(t, \mathbf{x}_{k-1}, \dots, \mathbf{x}_n) = \mathbf{x}_{k-1}$, $k \in \llbracket 2, i-1 \rrbracket$, $\mathbf{F}_i(t, \mathbf{x}_{i-1}, \dots, \mathbf{x}_n) =: \mathbf{x}_{i-1} + \bar{\mathbf{F}}_i(t, \mathbf{x}_j)$, $\mathbf{F}_k(t, \mathbf{x}) = \mathbf{x}_k$, k in $\llbracket i+1, j \rrbracket$ and $\mathbf{F}_\ell = \mathbf{0}$, ℓ in $\llbracket j+1, n \rrbracket$. In such a case, the noise added at the i^{th} level of the chain is only the $(i-1)^{\text{th}}$ iterated time integral of a Brownian motion

whose fluctuations are of order $t^{i-1/2}$ and the solution evolves according to its $(j-i)^{\text{th}}$ iterated integral in time. This simple feature illustrates the loss of fluctuations of the noise from level to level through its propagation to the chain and consequently in the corresponding spatial direction. This is the reason why the index of Hölder regularity increases when considering direction further and further away from the source of noise.

Taking then $\gamma = i - 1/2$ and $l = j - i$ in the above rule, we obtain that weak uniqueness may fail when $\tilde{\beta}_i^j < (2i - 3)/(2j - 1)$. Then, denoting for any i in $\llbracket 2, n \rrbracket$ and any j in $\llbracket i, n \rrbracket$, by $\tilde{\beta}_i^j$ the Hölder regularity index of the drift \mathbf{F}_i w.r.t. the j^{th} variable, we get that weak uniqueness may fail when $\tilde{\beta}_i^j < (2i - 3)/(2j - 1)$, which is precisely the result proved in [CdRM17]. Note that in this work the weak well-posedness is shown to hold as soon as the regularity indexes $(\tilde{\beta}_i^j)_{2 \leq i, j \leq n}$ satisfy $(2i - 3)/(2j - 1) < \tilde{\beta}_i^j \leq 1$ so that the result is *almost sharp*², in the sense that the critical case corresponding to an equality for the thresholds is still open.

Recall now that the thresholds obtained in our current work are $(2j - 2)/(2j - 1)$. Thus, even if it is clear that one cannot recover the thresholds of the regularity index obtained in the non degenerate framework, our thresholds remain greater than the ones obtained for the weak uniqueness in the same setting (so that they are then not covered by the class of counter-examples) and do not take into account the level of the chain (note indeed that the index “ i ”, which precisely relies on the level, does not appear in their definition).

Let us first forget this independence and, to have a better idea of what happened, let us come back to the simplest setting given above with $i = j$. This means that we are unable to prove strong well-posedness as soon as the corresponding Hölder regularity index β_j of the j^{th} component w.r.t. its j^{th} variable is lower than $(2j - 2)/(2j - 1) = 1 - 1/[2(j - 1/2)] = 1 - 1/[2(j - 1/2)]$ where $j - 1/2$ is precisely the self-similarity index of $(j - 1)^{\text{th}}$ iterated time integral of a Brownian motion and thus of the noise which perturbs the j^{th} component. Assuming (formally) some additional continuity on these thresholds we are hence not able to prove strong well-posedness for (1.2) as soon as $\alpha < 1 - 1/(2\gamma)$: we lose a factor 2 in comparison with the weak uniqueness threshold.

Unfortunately, we did not succeed in producing a corresponding class of counter-examples so that we can not claim that this gap is the price to pay to pass from weak to strong uniqueness. We nevertheless feel that this gap is intrinsic to the type of well-posedness asked on the system and we explain below how our thresholds appear and why they make sense for us. Especially, it will appear that they are *sharp w.r.t. the method employed* here (i.e. the Zvonkin Transform). We also emphasize that these thresholds can be also derived from the work of Catellier and Gubinelli [CG16] where non-degenerate fractional driven SDEs with singular drift are studied thanks to pathwise argument. They are also not able to prove strong well-posedness for regularity index α of the drift strictly lower than $1 - 1/(2H)$, H being the Hurst parameter of the fractional Brownian motion in $(0, 1]$. Our previous rule coincides with that threshold replacing H by the self-similarity index γ in $\{1/2, 3/2, \dots, n - 1/2\}$.

Before entering in the heuristic derivation of the thresholds, let us conclude by mentioning that there exists one dimensional results which clearly break the aforementioned (heuristic) rule, see e.g. the works of Gradinaru and Offret [GO13] or Bass and Chen [BC03]. Anyhow, the one dimensional case must be considered apart.

As already underlined, we tackle the strong well-posedness of (1.1) through Zvonkin Transform. Such an approach (which has also been used in the aforementioned non-degenerate results) relies on PDE techniques. It is indeed well known that the family of generators $(L_t)_{t \geq 0}$ of (1.1) is a family of linear partial differential operators of second order so that any SDE is connected with a PDE (the solution of the SDE being, in fact, the characteristic of the solution of the Cauchy problem associated with L). For instance, this connexion has been widely used in the last past five decades allowing to pass from regularization properties of PDE to well-posedness result for SDE and conversely (see e.g. the book of Bass [Bas97] for an account on weak well-posedness).

Keeping in mind this connexion, one may view the difference of critical Hölder indexes between strong and weak well posedness as a consequence of the different regularization properties required on the underlying PDE to prove the probabilistic result: roughly speaking the weak well posedness requires gradient bounds for the solution of the associated Cauchy problem while the strong well posedness requires in addition to control some of the cross derivatives of the solution.

Our approach to obtain such bounds is here based on perturbation techniques. For the weak uniqueness to hold, we are hence led to control the gradients of the Green kernel associated with a degenerate Ornstein-Uhlenbeck like process and an associated perturbation term involving the difference of the initial generator of (1.1) and the one of the Ornstein-Uhlenbeck proxy. For the strong uniqueness to hold, the quantities of

²the drift \mathbf{F}_1 may be assumed to be only in $L_q - L_p$ space for any $n^2 d/p + 2/q < 1$, $p \geq 2, q > 2$ or Hölder continuous for any strictly positive Hölder index.

interest are the derivatives of the gradient w.r.t. the non-degenerate variables for the Green kernel and the perturbative term.

Since the regularization properties of the Green kernel are the same in both cases, this additional control explains why the critical Hölder indexes increase for the strong uniqueness to hold.

The reason why the Hölder regularity index is blind to the level of the chain is a bit more involved at this stage of the presentation and we refer to subsections 1.3 and 1.4 below for a more detailed discussion on the subject. Let us only mention that in the weak case, the underlying Cauchy problem is investigated with a global source term f of independent regularity while in the strong case the source term should be *any* component of the drift of (1.1) itself. This is a consequence of the initial Zvonkin approach. Each of the $(\mathbf{F}_i)_{i \in \llbracket 1, n \rrbracket}$ has to be considered as a source term for the corresponding PDE.

Assuming that $d = 1$ and $n = 2$, using a perturbative approach to obtain gradient bounds w.r.t. the degenerate component on the considered Green kernel led to investigate the behavior of the ratio $|x_2|^{\tilde{\beta}_2^2}/t^{3/2}$ where x_2 is a spatial variable which has the same typical order as the integral of the Brownian motion³. Namely $|x_2| \asymp t^{3/2}$. Hence, $|x_2|^{\tilde{\beta}_2^2}/t^{3/2} \asymp t^{-3/2(1-\tilde{\beta}_2^2)}$. For weak uniqueness to hold, the point is then to impose that the previous ratio gives an integrable singularity in time. This precisely leads to the condition $-3/2(1-\tilde{\beta}_2^2) > -1 \Leftrightarrow \tilde{\beta}_2^2 > 1/3$.

The weak uniqueness thresholds for the full chain were derived similarly in [CdRM17] studying for $i \in \llbracket 2, n \rrbracket$, $|x_j|^{\tilde{\beta}_i^j}/t^{i-1/2}$, $j \in \llbracket i, n \rrbracket$, where x_j is a spatial variable which has the same typical order as the $(j-1)$ th iterated integral of the Brownian motion. Namely $|x_j| \asymp t^{j-1/2}$. The corresponding ratio thus satisfies $|x_j|^{\tilde{\beta}_i^j}/t^{i-1/2} \asymp t^{-(i-1/2)+\tilde{\beta}_i^j(j-1/2)}$. In order to obtain an integrable singularity one must take $-(i-1/2)+\tilde{\beta}_i^j(j-1/2) > -1 \Leftrightarrow \tilde{\beta}_i^j > (2i-3)/(2j-1)$ which is precisely the indicated threshold. Of course, the previous ratios appear through rather lengthy and technical perturbative procedures: the forward parametrix expansion in [CdR18] and the backward one in [CdRM17] (which allows to improve the Hölder thresholds on \mathbf{F}_1 w.r.t. [CdR18]).

Recall now that, to obtain the strong well-posedness result for the chain (1.1), the quantities of interest are the derivatives of the gradient w.r.t. the non-degenerate variables for the considered Green kernel. This additional differentiation leads to consider the ratios $|\mathbf{x}_j|^{\beta_j}/t^{(j-1/2)+1/2}$, $j \in \llbracket 1, n \rrbracket$ i.e. the time singularity is higher of degree 1/2: this precisely gives the factor 2 lost when passing from weak to strong uniqueness. Indeed, considering again the associated typical scale for the variable \mathbf{x}_j , which is $t^{j-1/2}$, we write $|\mathbf{x}_j|^{\beta_j}/t^{(j-1/2)+1/2} \asymp t^{-j+\beta_j(j-1/2)}$. To obtain an integrable singularity we get $-j+\beta_j(j-1/2) > -1 \Leftrightarrow \beta_j > (2j-2)/(2j-1)$, which is the indicated threshold.

These features, Zvonkin transform and perturbative analysis are detailed respectively in Section 1.3 and 1.4 below.

1.2 Notations, assumptions and main result

Some notations. We will denote by a bold letter \mathbf{x}, \mathbf{y} any element of \mathbb{R}^{nd} , writing as well $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ where for $i \in \llbracket 2, n \rrbracket$, $\mathbf{x}_i \in \mathbb{R}^d$. For practical purpose we will be led in our analysis to consider *subcomponents* of a vector $\mathbf{x} \in \mathbb{R}^{nd}$. Namely, for any $0 \leq i \leq j \leq n$ and $\mathbf{x} \in \mathbb{R}^{nd}$, we introduce the notation $\mathbf{x}_{i:j} := (\mathbf{x}_i, \dots, \mathbf{x}_j)$. Accordingly, we write the drift as the mapping

$$\begin{aligned} (s, \mathbf{x}) \in \mathbb{R}^+ \times \mathbb{R}^{nd} &\mapsto \mathbf{F}(s, \mathbf{x}) = (\mathbf{F}_1(s, \mathbf{x}), \dots, \mathbf{F}_n(s, \mathbf{x})) \\ &= (\mathbf{F}_1(s, \mathbf{x}), \mathbf{F}_2(s, \mathbf{x}), \mathbf{F}_3(s, \mathbf{x}_{2:n}) \cdots, \mathbf{F}_n(s, \mathbf{x}_{n-1:n})), \end{aligned}$$

from the specific structure of the drift appearing in (1.1).

For $f \in C^1(\mathbb{R}^{nd}, \mathbb{R}^k)$, $k \in \{1, d\}$, we denote for all $i \in \llbracket 1, n \rrbracket$, by $D_{\mathbf{x}_i} f(\mathbf{x})$ the Jacobian matrix of the derivative of f w.r.t. to its \mathbb{R}^d -valued variable \mathbf{x}_i . As shortened form, and when no ambiguity is possible, we also write for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{nd}$, $D_{\mathbf{x}_i} f(\mathbf{x}) = D_i f(\mathbf{x})$ and $D_{\mathbf{y}_i} f(\mathbf{y}) = D_i f(\mathbf{y})$. Also, if $k = 1$ we denote by $\mathbf{D}f(\mathbf{x}) = (D_1 f(\mathbf{x}) \cdots D_n f(\mathbf{x}))^*$ the full gradient of the function f at point \mathbf{x} .

Let $f : \mathbb{R}^{nd} \rightarrow \mathbb{R}^k$ and $\beta := (\beta_1, \dots, \beta_n) \in (0, 1]^n$ be a multi-index. We say that f is uniformly β -Hölder continuous if for all $j \in \llbracket 1, n \rrbracket$

$$[f_j]_{\beta_j} := \sup_{(\mathbf{z}_{1:j-1}, \mathbf{z}_{j+1:n}) \in \mathbb{R}^{(n-1)d}, z \neq z', (z, z') \in (\mathbb{R}^d)^2} \frac{|f(s, \mathbf{z}_{1:j-1}, z, \mathbf{z}_{j+1:n}) - f(s, \mathbf{z}_{1:j-1}, z', \mathbf{z}_{j+1:n})|}{|z - z'|^{\beta_j}} < +\infty. \quad (1.3)$$

³Here $t^{3/2}$ is precisely the singularity coming from the differentiation of the density of the degenerate Ornstein-Uhlenbeck process w.r.t. the degenerate variable and $|x_2|^{\tilde{\beta}_2^2}$ typically reflects the behavior of a source term with Hölder regularity index $\tilde{\beta}_2^2$ w.r.t. the degenerate variable.

For a smooth function $\Psi : [0, T] \times \mathbb{R}^{nd} \rightarrow \mathbb{R}^{nd}$, where $T > 0$ is a fixed given time, writing for $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$, $\Psi(t, \mathbf{x}) = (\Psi_1(t, \mathbf{x}), \dots, \Psi_n(t, \mathbf{x}))^*$ where for each $i \in \llbracket 1, n \rrbracket$, Ψ_i is \mathbb{R}^d valued, we denote by

$$\|\mathbf{D}\Psi\|_\infty := \sum_{i=1}^n \sup_{(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}} \|\mathbf{D}\Psi_i(t, \mathbf{x})\|, \|\mathbf{D}(D_1\Psi)\|_\infty := \sum_{i=1}^n \sup_{(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}} \|\mathbf{D}(D_1\Psi_i)(t, \mathbf{x})\|, \quad (1.4)$$

where in the above equation $\|\cdot\|$ stands for a tensor norm in the appropriate corresponding dimension. Precisely, $\mathbf{D}\Psi_i(t, \mathbf{x}) \in \mathbb{R}^{nd} \otimes \mathbb{R}^d$ and $\mathbf{D}(D_1\Psi_i)(t, \mathbf{x}) \in \mathbb{R}^{nd} \otimes \mathbb{R}^d \otimes \mathbb{R}^d$.

Assumptions We will assume throughout the paper that the following conditions hold.

(ML) The coefficients \mathbf{F} and σ are measurable in time and $\mathbf{F}(t, \mathbf{0})$ is bounded. Also, the diffusion coefficient σ is uniformly Lipschitz continuous in space, uniformly in time, i.e. there exists $\kappa > 0$ s.t. for all $t \geq 0$, $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$:

$$|\sigma(t, \mathbf{x}) - \sigma(t, \mathbf{x}')| \leq \kappa |\mathbf{x} - \mathbf{x}'|.$$

(UE) The diffusion matrix $a := \sigma\sigma^*$ is uniformly elliptic and bounded, uniformly in time, i.e. there exists $\Lambda \geq 1$ s.t. for all $t \geq 0$, $(\mathbf{x}, \zeta) \in \mathbb{R}^{nd} \times \mathbb{R}^d$:

$$\Lambda^{-1}|\zeta|^2 \leq \langle a(t, \mathbf{x})\zeta, \zeta \rangle \leq \Lambda|\zeta|^2.$$

(T $_\beta$) For all $j \in \llbracket 1, n \rrbracket$, the functions $(\mathbf{F}_i)_{i \in \llbracket 1, (j+1) \wedge n \rrbracket}$ are uniformly β_j -Hölder continuous in the j^{th} spatial variable with $\beta_j \in (\frac{2j-2}{2j-1}, 1]$, uniformly w.r.t. the other spatial variables of \mathbf{F}_i and in time. In particular, there exists a finite constant C_β s.t.

$$\max_{(i, j) \in \llbracket 1, n \rrbracket^2} \sup_{s \in [0, T]} [(\mathbf{F}_i)_j(s, \cdot)]_{\beta_j} \leq C_\beta.$$

(H $_\eta$) For all $i \in \llbracket 2, n \rrbracket$, there exists a closed convex subset $\mathcal{E}_{i-1} \subset GL_d(\mathbb{R})$ (set of invertible $d \times d$ matrices) s.t., for all $t \geq 0$ and $(\mathbf{x}_{i-1}, \dots, \mathbf{x}_n) \in \mathbb{R}^{(n-i+2)d}$, $D_{\mathbf{x}_{i-1}}F_i(t, \mathbf{x}_{i-1}, \dots, \mathbf{x}_n) \in \mathcal{E}_{i-1}$. For example, \mathcal{E}_{i-1} may be a closed ball included in $GL_d(\mathbb{R})$, which is an open set. Moreover, $D_{\mathbf{x}_{i-1}}\mathbf{F}_i$ is η -Hölder continuous w.r.t. \mathbf{x}_{i-1} uniformly in $\mathbf{x}_{i:n}$ and time. We also assume without loss of generality that $\eta \in (0, \inf_{j \in \llbracket 2, n \rrbracket} \{\beta_j - \frac{2j-2}{2j-1}\})$, i.e. η is meant to be *small*.

From now on, we will say that assumption **(A)** is in force provided that **(ML)**, **(UE)**, **(T $_\beta$)**, **(H $_\eta$)** hold.

Main result. The main result of this work is the following theorem.

Theorem 1 (Strong uniqueness for the degenerate system (1.1)). *Under **(A)** there exists a unique strong solution to system (1.1).*

Remark 1. *Still in comparison with the results obtained in the non-degenerate cases, and especially the one of Krylov and Röckner [KR05], we do not tackle the case of drift in $L_q - L_p$ w.r.t. the first (and then non-degenerate) variable. This is only to keep our result as clear as possible and to concentrate on the novelty of the approach we use here. We are anyhow confident that these specific drifts could be handled. Indeed, all the intermediate results needed to perform the analysis in that setting seem to be already available. We refer to subsection 1.4 for further details.*

1.3 Proof of the main result: Zvonkin Transform and smoothing properties of the PDE associated with (1.1)

We emphasize that under our assumptions, it follows from [CdRM17] that (1.1) is well posed in the weak sense. Hence, from Yamada-Watanabe theorem it is sufficient to prove that strong (or pathwise) uniqueness holds to prove strong well posedness. To do so, our main strategy rests upon the Zvonkin transform initiated by Zvonkin in [Zvo74] which has been widely used during the last decade to prove strong well posedness, see *e.g.*, and heavily relies on the connexion between SDE and PDE as we already emphasized.

Introducing the embedding matrix B from \mathbb{R}^d into \mathbb{R}^{nd} , i.e. $B = (\mathbf{I}_{d,d}, \mathbf{0}_{d,d}, \dots, \mathbf{0}_{d,d})^* = (\mathbf{I}_{d,d}, \mathbf{0}_{d,(n-1)d})^*$, where “ $*$ ” stands for the transpose, we rewrite (1.1) in the shortened form

$$d\mathbf{X}_t = \mathbf{F}(t, \mathbf{X}_t)dt + B\sigma(t, \mathbf{X}_t)dW_t,$$

where $\mathbf{F} = (\mathbf{F}_1, \dots, \mathbf{F}_n)$ is an \mathbb{R}^{nd} -valued function. For all $\varphi \in C_0^2(\mathbb{R}^{nd}, \mathbb{R})$ and $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$ let

$$L_t \varphi(\mathbf{x}) = \langle \mathbf{F}(t, \mathbf{x}), \mathbf{D}\varphi(t, \mathbf{x}) \rangle + \frac{1}{2} \text{Tr} \left(a(t, \mathbf{x}) D_{\mathbf{x}_1}^2 \varphi(\mathbf{x}) \right), \quad (1.5)$$

where $a = \sigma \sigma^*$ denotes the generator associated with (1.1). We then *formally* associate the SDE (1.1) with the following systems of PDEs:

$$\begin{cases} (\partial_t \mathbf{u}_i + L_t \mathbf{u}_i)(t, \mathbf{x}) = \mathbf{F}_i(t, \mathbf{x}), & (t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}, \\ \mathbf{u}_i(T, \mathbf{x}) = \mathbf{0}_d, & i \in \llbracket 1, n \rrbracket, \end{cases} \quad (1.6)$$

and we denote by $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ its global solution. Let now $(\mathbf{F}^m)_{m \geq 0}, (a^m)_{m \geq 0}$ denote two sequences of mollified coefficients satisfying assumption **(A)** uniformly in m that are infinitely differentiable functions with bounded derivatives of all, order greater than 1 for \mathbf{F}^m , and converging in supremum norm to (\mathbf{F}, a) (such sequences are easily obtained from [CdR17]). Then, for each m , the regularized systems of PDEs associated with (1.6) write:

$$\begin{cases} (\partial_t \mathbf{u}_i^m + L_t^m \mathbf{u}_i^m)(t, \mathbf{x}) = \mathbf{F}_i^m(t, \mathbf{x}), & (t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}, \\ \mathbf{u}_i^m(T, \mathbf{x}) = \mathbf{0}_d, & i \in \llbracket 1, n \rrbracket, \end{cases} \quad (1.7)$$

where L_t^m is obtained from (1.5) replacing \mathbf{F} by \mathbf{F}^m and a by a^m .

The above system (1.7) is well posed and admits a unique smooth solution $\mathbf{U}^m = (\mathbf{u}_1^m, \dots, \mathbf{u}_n^m)$. Hence, applying Itô's Formula, one easily deduces that

$$\int_0^t \mathbf{F}(s, \mathbf{X}_s) ds = -\mathbf{U}^m(0, \mathbf{x}) + \mathbf{U}^m(t, \mathbf{X}_t) - \int_0^t \mathbf{D}\mathbf{U}^m(s, \mathbf{X}_s) B \sigma(s, \mathbf{X}_s) dW_s + \mathcal{R}_t^m(\mathbf{X}), \quad (1.8)$$

where

$$\mathcal{R}_t^m(\mathbf{X}) := \int_0^t \mathbf{F}(s, \mathbf{X}_s) - \mathbf{F}^m(s, \mathbf{X}_s) ds - \int_0^t (L_s - L_s^m) \mathbf{U}^m(s, \mathbf{X}_s) ds.$$

The representation (1.8) is what is usually called the *Zvonkin Transform*. The main thing in this representation is that the *bad drift* can be, up to a remainder, rewritten in term of the solution of the system of PDEs (1.7) for which the source term is precisely a mollified version of the *bad drift* itself.

Then, the main idea consists in taking advantage of the regularization properties of the operator L^m (uniformly in m) and expect that the solutions \mathbf{U}^m , $m \geq 0$ will be smoother than the source term \mathbf{F} so that the right hand side of (1.8) is smoother than the integrand of the left hand side of the considered equation. In other words, we are looking for a *good regularization theory* for the PDE (1.7) uniformly on the mollification argument. This *good regularization theory* is summarized in the following crucial result whose proof is, in fact, the main subject of this work and is postponed to Section 2.

Theorem 2. *For $T > 0$ small enough⁴, there exists a constant $C_T := C_T(\mathbf{A}) > 0$ satisfying $C_T \rightarrow 0$ when $T \rightarrow 0$ such that for every $m \geq 0$, the solution \mathbf{U}^m satisfies with the notation of (1.4):*

$$\|\mathbf{D}\mathbf{U}^m\|_\infty + \|\mathbf{D}(D_1 \mathbf{U}^m)\|_\infty \leq C_T. \quad (1.9)$$

Remark 2 (On well-posedness of the initial PDE (1.6)). *We also point out that, from the uniformity in m in the previous theorem, we could also derive some regularizing properties for the system (1.6) through appropriate compactness arguments. Indeed, as it will appear in the proof of this result, we are in fact able to control uniformly the Hölder moduli of the gradients and of the second order derivatives w.r.t. the non-degenerate direction (see Lemma 10 and Remark 4). These controls precisely allow to derive, through Arzelà-Ascoli argument, a well-posedness result for equation (1.6) under the sole assumption **(A)** as well as the above gradient estimates.*

Let now \mathbf{X} and \mathbf{X}' be two solutions of (1.1). Using the representation (1.8) to express the difference of the *bad drift* in terms of the function \mathbf{U}^m and its derivative, we write:

$$\begin{aligned} & \mathbf{X}_t - \mathbf{X}'_t \\ = & \mathbf{U}^m(t, \mathbf{X}_t) - \mathbf{U}^m(t, \mathbf{X}'_t) - \int_0^t [\mathbf{D}\mathbf{U}^m(s, \mathbf{X}_s) B \sigma(s, \mathbf{X}_s) - \mathbf{D}\mathbf{U}^m(s, \mathbf{X}'_s) B \sigma(s, \mathbf{X}'_s)] dW_s \\ & + \left[\mathcal{R}_t^m(\mathbf{X}) - \mathcal{R}_t^m(\mathbf{X}') \right] + \int_0^t B [\sigma(s, \mathbf{X}_s) - \sigma(s, \mathbf{X}'_s)] dW_s. \end{aligned}$$

⁴By “small enough” we mean that there exists a time $\mathcal{T} > 0$ depending on known parameters in **(A)** s.t. for all $T \leq \mathcal{T}$ the statement of the theorem holds.

Take then the supremum in time of the square of the difference. Passing to the expectation, a convexity inequality then leads to the following estimate:

$$\begin{aligned} \mathbb{E} \left[\sup_{t \leq T} |\mathbf{X}_t - \mathbf{X}'_t|^2 \right] &\leq 5 \left(\mathbb{E} \left[\sup_{t \leq T} |\mathbf{U}^m(t, \mathbf{X}_t) - \mathbf{U}^m(t, \mathbf{X}'_t)|^2 \right] \right. \\ &\quad + \mathbb{E} \left[\int_0^T \| [\mathbf{D}\mathbf{U}^m B](s, \mathbf{X}_s) - [\mathbf{D}\mathbf{U}^m B](s, \mathbf{X}'_s) \|^2 \|\sigma\|_\infty^2 ds \right] \\ &\quad \left. + \mathbb{E} \left[\int_0^T (\|\mathbf{D}\mathbf{U}^m B\|_\infty + 1) \|\sigma(s, \mathbf{X}_s) - \sigma(s, \mathbf{X}'_s)\|^2 ds \right] + 2\|\mathcal{R}^m(\cdot)\|_\infty^2 \right). \end{aligned}$$

Note that thanks to the particular structure of B one has $\mathbf{D}\mathbf{U}^m B = (D_{\mathbf{x}_1} \mathbf{U}^m, \mathbf{0}_{d,d}, \dots, \mathbf{0}_{d,d})^*$. Hence, thanks to Theorem 2 and Grönwall's lemma, there exists $\bar{C}_T := \bar{C}_T(C_T, \sigma, n, d, T)$ satisfying $\bar{C}_T \rightarrow 0$ when T goes to 0 such that

$$\mathbb{E} \left[\sup_{t \leq T} |\mathbf{X}_t - \mathbf{X}'_t|^2 \right] \leq \bar{C}_T \mathbb{E} \left[\sup_{t \leq T} |\mathbf{X}_t - \mathbf{X}'_t|^2 \right] + 10\|\mathcal{R}^m(\cdot)\|_\infty^2. \quad (1.10)$$

Letting $m \rightarrow +\infty$ and choosing T small enough so that $\bar{C}_T \leq 1/2$, we deduce that strong uniqueness holds on a sufficiently small time interval. Iterating this procedure in time gives the result on \mathbb{R}^+ from usual Markov arguments involving the regular versions of conditional expectations, see e.g. [SV79].

1.4 Regularization properties of the underlying PDE (1.6): strategy of proof and primer

As mentioned above, the regularization properties of the PDE (1.7) given by estimate (1.9) in Theorem 2 are the core of this work. Smoothing properties of linear partial differential operators of second order with non-degenerate diffusion matrix have been widely studied in the literature and, in that setting, the estimates of Theorem 2 are well known (see *e.g.* the book of Friedman [Fri64] or of Bass [Bas97]). In our case, the story is rather different since the diffusion matrix Ba of the system is totally degenerate in the directions 2 to n . However, as we already emphasized, the non-degeneracy condition assumed on the family of Jacobians $(D_{\mathbf{x}_{i-1}} \mathbf{F}_i)_{i \in \llbracket 2, n \rrbracket}$ allows the noise to propagate in the mentioned directions thanks to the drift. It can be viewed as a weak type of Hörmander condition. Under such a condition, the operator L^m with mollified coefficients is said to be hypoelliptic⁵ and it is well known that hypoelliptic differential operators also have some smoothing properties (see the seminal work of Hörmander [Hör67] or, for a probabilistic viewpoint, the ouvrage of Stroock [Str08]). The tricky point in our *weak Hörmander setting* is that the pointwise gradient estimates (1.9) of Theorem 2 had, to the best of our knowledge, not been established yet. Although such a setting has already been considered by several authors (see *e.g.* Delarue and Menozzi [DM10] for density estimates, Menozzi [Men11], [Men18] and Priola [Pri15] for the martingale problem and also Bramanti, Cupini, Lanconelli and Priola [BCLP10], [BCLP13] for related L^p estimates and Bramanti and Zhu [BZ11] for the VMO framework). We can mention the work of Lorenzi [Lor05] which gives gradient estimates in the degenerate kinetic like case ($n = 2$ in our framework) when the diffusion coefficient is sufficiently smooth and the drift linear. We point out that our main estimate in Theorem 2 needs precisely to be uniform w.r.t. the mollification parameter and therefore does not depend on the smoothness of \mathbf{F}^m, a^m , but only on known parameters appearing in (\mathbf{A}) . Again, this is what would also allow to transfer those bounds to equation (1.6) from a suitable compactness argument, extending well known results for non-degenerate diffusions with Hölder coefficients to the current degenerate setting.

To prove this result our main strategy rests upon the *parametrix approach* see *e.g.* the work of McKean and Singer [MS67] or the book of Friedman [Fri64]. Roughly speaking, it consists in a perturbation argument of the operator L^m which is expanded around a *good proxy*, usually denoted by \tilde{L}^m (we keep here the super-scripts in m to emphasize that the perturbative technique we perform will concern the system (1.7) with mollified coefficients). The terminology of *good* in our setting relies to the fact that the operator \tilde{L}^m is the generator of the “closest” Gaussian approximation $\tilde{\mathbf{X}}^m$ of \mathbf{X}^m which has generator L^m . In our case, such a process is well known and is the linearized (with respect to the source of noise) version of (1.1) whose coefficients are frozen along the curve $(\boldsymbol{\theta}_{s,t}^m)_{s \in [t, T]}$ that solves the deterministic counterpart of (1.1) with mollified coefficients (*i.e.*

⁵Pay attention that this is not the case for L whose coefficients do not have the required smoothness in (\mathbf{T}_β) to compute the corresponding Lie brackets.

with $\sigma^m \equiv \mathbf{0}_{d,d}$ namely, $\dot{\theta}_{s,t}^m = \mathbf{F}^m(s, \theta_{s,t}^m)$. This process may be seen as a (non-linear) generalization of the so-called Kolmogorov example [Kol34] and we refer the reader to the work of Delarue and Menozzi [DM10] and Menozzi [Men11] for more explanations. Having this proxy at hand, the parametrix procedure consists in deriving the desired estimates for the proxy and control the expansion error.

In [CdRM17], Chaudru de Raynal and Menozzi successfully used this approach in its *backward* form to prove weak well posedness of (1.1) under less restrictive assumptions (the critical thresholds for the Hölder exponents being smaller as indicated above). In that case, the curve along which the system is frozen for the proxy is the solution of the *backward* deterministic counterpart of (1.1). This *backward* approach is very suitable when investigating the martingale problem associated with our main system since it allows to control subtly the expansion error associating precisely the coefficients \mathbf{F}_i with their corresponding differentiation operator D_i and does not require any mollification of the coefficients. Unfortunately, when trying to obtain estimates on the derivatives of the solutions of the PDE, the *backward* approach is not appropriate since the corresponding proxy does not provide an exact density and this fact does not allow to benefit from cancellation techniques which are very helpful in this context (see paragraph below).

Hence, our parametrix approach will be of *forward*⁶ form as done in the work of Chaudru de Raynal [CdR17]. This is, in fact, a non-trivial generalization of the approach developed in the aforementioned paper where the strong well posedness of (1.1) is obtained when $n = 2$. Indeed, the strategy used in [CdR17] is not adapted to this general case because of some subtle phenomena appearing only when $n \geq 3$. In particular, the singularities appearing when considering the remainder term of the parametrix were in [CdR17] equilibrated *at hand* through elementary cancellation arguments, whereas the current approach takes advantage of the full-force duality results between Besov spaces (see Sections 1.4.2 and 2.2 below). This forward perturbative approach has also been successfully used in [CdRHM18] to establish some weaker regularization properties of the PDE (1.6) through appropriate Schauder estimates.

1.4.1 Regularizing properties of the degenerate Ornstein-Uhlenbeck proxy

When exploiting such a *forward* parametrix approach, a good primer to understand what could be, at best, expected, consists in investigating the regularization property of the proxy operator \tilde{L} . To be as succinct as possible, let us consider the case where $(\tilde{L}_t)_{t \geq 0}$ is the generator of a degenerate Ornstein-Uhlenbeck process $(\tilde{\mathbf{X}}_t)_{t \geq 0}$ with dynamics:

$$d\tilde{\mathbf{X}}_t = \mathbf{A}_t \tilde{\mathbf{X}}_t dt + BdW_t, \quad (1.11)$$

where \mathbf{A}_t is the $nd \times nd$ matrix with sub-diagonal block $\mathbf{a}_{i,i-1}(t)$ of size $d \times d$ and $\mathbf{0}_{d,d}$ elsewhere. In particular,

$$\mathbf{A}_t = \begin{pmatrix} \mathbf{0}_{d,d} & \cdots & \cdots & \cdots & \mathbf{0}_{d,d} \\ \mathbf{a}_{2,1}(t) & \mathbf{0}_{d,d} & \cdots & \cdots & \mathbf{0}_{d,d} \\ \mathbf{0}_{d,d} & \mathbf{a}_{3,2}(t) & \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} \\ \vdots & \mathbf{0}_{d,d} & \ddots & \vdots & \\ \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} & \mathbf{a}_{n,n-1}(t) & \mathbf{0}_{d,d} \end{pmatrix}. \quad (1.12)$$

The entries $(\mathbf{a}_{i,i-1}(t))_{i \in [2,n]}$ are uniformly in time non-degenerate elements of $\mathbb{R}^d \otimes \mathbb{R}^d$ (which expresses the weak Hörmander condition). The corresponding generator \tilde{L}_t writes for all $\varphi \in C_0^2(\mathbb{R}^{nd}, \mathbb{R})$:

$$\tilde{L}_t \varphi(\mathbf{x}) = \langle \mathbf{A}_t \mathbf{x}, \mathbf{D}\varphi(\mathbf{x}) \rangle + \frac{1}{2} \Delta_{\mathbf{x}_1} \varphi(\mathbf{x}).$$

In such a case, each component $\tilde{\mathbf{u}}_i$, $i \in [1, n]$ of the solution $\tilde{\mathbf{U}}$ of the corresponding system of PDEs

$$\begin{cases} (\partial_t + \tilde{L}_t) \tilde{\mathbf{U}}(t, \mathbf{x}) = \mathbf{F}(t, \mathbf{x}), & (t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}, \\ \tilde{\mathbf{U}}(T, \mathbf{x}) = \mathbf{0}_{nd}, \end{cases} \quad (1.13)$$

where \mathbf{F} is a non-linear (non-mollified) source satisfying (\mathbf{T}_β) , writes

$$\tilde{\mathbf{u}}_i(t, \mathbf{x}) = - \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \mathbf{F}_i(s, \mathbf{y}) \tilde{p}(t, s, \mathbf{x}, \mathbf{y}). \quad (1.14)$$

⁶Meaning that the freezing curve θ solves the corresponding ODE associated with (1.1) in a forward form.

Above, \tilde{p} stands for the transition density of the Gaussian process $(\mathbf{X}_v)_{v \geq 0}$ with dynamics (1.11). Using the resolvent associated with $(\mathbf{A}_v)_{v \in [t, s]}$, i.e. $\partial_s \tilde{\mathbf{R}}_{s,t} = \mathbf{A}_s \tilde{\mathbf{R}}_{s,t}$, $\tilde{\mathbf{R}}_{t,t} = \mathbf{I}_{nd, nd}$, the above equation can be explicitly integrated. Precisely, for a fixed starting point \mathbf{x} at time t :

$$\tilde{\mathbf{X}}_v = \tilde{\mathbf{R}}_{v,t} \mathbf{x} + \int_t^v \tilde{\mathbf{R}}_{v,u} B dW_u. \quad (1.15)$$

Hence, the covariance between given times $t < s$ writes $\tilde{\mathbf{K}}_{s,t} := \int_t^s \tilde{\mathbf{R}}_{s,u} B B^* \tilde{\mathbf{R}}_{s,u}^* du$. From (1.15), the density at time $v = s$ and at the spatial point \mathbf{y} therefore writes:

$$\tilde{p}(t, s, \mathbf{x}, \mathbf{y}) = \frac{1}{(2\pi)^{\frac{nd}{2}} \det(\tilde{\mathbf{K}}_{s-t})^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \left\langle (\tilde{\mathbf{K}}_{s-t})^{-1} (\tilde{\mathbf{R}}_{s,t} \mathbf{x} - \mathbf{y}), \tilde{\mathbf{R}}_{s,t} \mathbf{x} - \mathbf{y} \right\rangle\right). \quad (1.16)$$

Note that the resolvent also appears in (1.15) and in the density. Since the drift in (1.11) is unbounded, the term $\tilde{\mathbf{R}}_{s,t} \mathbf{x}$ actually corresponds to the transport of the initial condition \mathbf{x} through the associated deterministic and linear differential system. It is well known, see e.g. [DM10] and Section 2.1 below, that the covariance $\tilde{\mathbf{K}}_{s,t}$ enjoys what we will call a *good scaling property*. Precisely, for a given $T > 0$ there exists $C := C((\mathbf{A}_v)_{v \in [0, T]}, T) \geq 1$ s.t. for all $\boldsymbol{\xi} \in \mathbb{R}^{nd}$,

$$C^{-1}(s-t)^{-1} |\mathbb{T}_{s-t} \boldsymbol{\xi}|^2 \leq \langle \tilde{\mathbf{K}}_{s,t} \boldsymbol{\xi}, \boldsymbol{\xi} \rangle \leq C(s-t)^{-1} |\mathbb{T}_{s-t} \boldsymbol{\xi}|^2, \quad (1.17)$$

where for all $u > 0$, we denote by \mathbb{T}_u the *intrinsic* scale matrix:

$$\mathbb{T}_u = \begin{pmatrix} u \mathbf{I}_{d,d} & \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} \\ \mathbf{0}_{d,d} & u^2 \mathbf{I}_{d,d} & \mathbf{0}_{d,d} & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} & u^n \mathbf{I}_{d,d} \end{pmatrix}. \quad (1.18)$$

Importantly, the *good scaling property* stated in (1.17) indicates that, for a given initial time t and for all $i \in \llbracket 1, n \rrbracket$, each \mathbb{R}^d -valued component $\tilde{\mathbf{X}}_s^i$ has typical fluctuations of order $(s-t)^{i-1/2}$ which corresponds to those of the $(i-1)^{\text{th}}$ iterated integrals of the Brownian motion. Accordingly, we derive that the frozen density \tilde{p} also satisfies the bound

$$\tilde{p}(t, s, \mathbf{x}, \mathbf{y}) \leq \frac{C}{(s-t)^{\frac{nd}{2}}} \exp\left(-C^{-1}(s-t) |\mathbb{T}_{s-t}^{-1} (\tilde{\mathbf{R}}_{s,t} \mathbf{x} - \mathbf{y})|^2\right) =: C \hat{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}),$$

where, up to a modification of the constants involved, $\int_{\mathbb{R}^{nd}} d\mathbf{y} \hat{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}) = 1$. Similarly, the derivatives of \tilde{p} will be bounded by a density of the form $\hat{p}_{C^{-1}}$ up to an additional multiplicative contribution reflecting the time-singularities associated with the differentiation index. Precisely, there exists \bar{C} s.t. for any $l \in \llbracket 1, n \rrbracket$, $r \in \{0, 1\}$:

$$\begin{aligned} |D_{\mathbf{x}_l} D_{\mathbf{x}_1}^r \tilde{p}(t, s, \mathbf{x}, \mathbf{y})| &\leq \frac{\bar{C}}{(s-t)^{\frac{nd}{2} + (l-\frac{1}{2}) + \frac{r}{2}}} \exp\left(-C^{-1}(s-t) |\mathbb{T}_{s-t}^{-1} (\tilde{\mathbf{R}}_{s,t} \mathbf{x} - \mathbf{y})|^2\right) \\ &\leq \frac{\bar{C}}{(s-t)^{(l-\frac{1}{2}) + \frac{r}{2}}} \hat{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}). \end{aligned} \quad (1.19)$$

We refer to the proof of Proposition 5 for a complete version of this statement.

To prove estimate (1.9) of Theorem 2 for the current system (1.13), it follows from the specific structure of the matrix B that we have to estimate for any $l \in \llbracket 1, n \rrbracket$ the quantities $D_{\mathbf{x}_l} D_{\mathbf{x}_1}^r \tilde{\mathbf{u}}_i(t, \mathbf{x})$, $r \in \{0, 1\}$. From (1.19), we thus have

$$|D_{\mathbf{x}_l} D_{\mathbf{x}_1}^r \tilde{\mathbf{u}}_i(t, \mathbf{x})| \leq \bar{C} \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} |\mathbf{F}_i(s, \mathbf{y})| (s-t)^{-(l-1/2) - r/2} \hat{p}_{C^{-1}}(t, s, \mathbf{x}, \mathbf{y}). \quad (1.20)$$

We now face two problems: first the \mathbf{F}_i are unbounded, second the above time singularity is, as is, not integrable. Let us consider the worst case *i.e.* when $r = 1$. To smoothen the time singularity, the main idea consists in using the regularity of the source term \mathbf{F}_i by exploiting precisely the fact that, once integrated through the variables \mathbf{y}_l to \mathbf{y}_n , the transition density \tilde{p} does not depend on the variable \mathbf{x}_l anymore. This is due to the structure of \mathbf{A} in (1.12), which in particular yields that the resolvent $(\tilde{\mathbf{R}}_{s,t})_{0 \leq t \leq s \leq T}$ is lower triangular. It is

hence equal to 0 when we differentiate it w.r.t. \mathbf{x}_l : this is what will be called a *cancellation (or centering) argument* in the following. Precisely, denoting for conciseness by $\boldsymbol{\theta}_{s,t}(\mathbf{x}) = \mathbf{R}_{s,t}\mathbf{x}$ (which is coherent with the notation below when handling non-linear flows), we write:

$$\int_{\mathbb{R}^{(l-1)d}} d\mathbf{y}_{1:l-1} \mathbf{F}_i(t, \mathbf{y}_1, \dots, \mathbf{y}_{l-1}, \boldsymbol{\theta}_{s,t}^l(\mathbf{x}), \dots, \boldsymbol{\theta}_{s,t}^n(\mathbf{x})) D_{\mathbf{x}_1} D_{\mathbf{x}_l} \int_{\mathbb{R}^{(n-(l-1))d}} d\mathbf{y}_{l:n} \tilde{p}(t, s, \mathbf{x}, \mathbf{y}) = 0.$$

When using this property, we obtain that

$$\begin{aligned} & |D_{\mathbf{x}_1} D_{\mathbf{x}_l} \tilde{\mathbf{u}}_i(t, \mathbf{x})| \\ &= \left| \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \left(\mathbf{F}_i(s, \mathbf{y}) - \mathbf{F}_i(t, \mathbf{y}_1, \dots, \mathbf{y}_{l-1}, \boldsymbol{\theta}_{s,t}^l(\mathbf{x}), \dots, \boldsymbol{\theta}_{s,t}^n(\mathbf{x})) \right) D_{\mathbf{x}_1} D_{\mathbf{x}_l} \tilde{p}(t, s, \mathbf{x}, \mathbf{y}) \right|. \end{aligned}$$

We thus obtain from (1.19):

$$\begin{aligned} & |D_{\mathbf{x}_1} D_{\mathbf{x}_l} \tilde{\mathbf{u}}_i(t, \mathbf{x})| \\ &\leq \bar{C} \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \left| \mathbf{F}_i(s, \mathbf{y}) - \mathbf{F}_i(t, \mathbf{y}_1, \dots, \mathbf{y}_{l-1}, \boldsymbol{\theta}_{s,t}^l(\mathbf{x}), \dots, \boldsymbol{\theta}_{s,t}^n(\mathbf{x})) \right| \\ &\quad \times (s-t)^{-(l-1/2)-1/2} \hat{p}_{C-1}(t, s, \mathbf{x}, \mathbf{y}). \end{aligned}$$

Then, using the regularity assumed of \mathbf{F}_i , which satisfies (\mathbf{T}_β) , we get that for some constant C (which possibly changes from line to line)

$$\begin{aligned} |D_{\mathbf{x}_1} D_{\mathbf{x}_l} \tilde{\mathbf{u}}_i(t, \mathbf{x})| &\leq C \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \sum_{j=l}^n |\mathbf{y}_j - \boldsymbol{\theta}_{s,t}^j(\mathbf{x})|^{\beta_j} (s-t)^{-(l-1/2)-1/2} \hat{p}_{C-1}(t, s, \mathbf{x}, \mathbf{y}) \\ &\leq C \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \sum_{j=l}^n \left| \frac{\mathbf{y}_j - \boldsymbol{\theta}_{s,t}^j(\mathbf{x})}{(s-t)^{j-1/2}} \right|^{\beta_j} (s-t)^{\beta_j(j-1/2)} \\ &\quad \times (s-t)^{-(l-1/2)-1/2} \hat{p}_{C-1}(t, s, \mathbf{x}, \mathbf{y}) \\ &\leq C \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \sum_{j=l}^n (s-t)^{-(l-1/2)-1/2+\beta_j(j-1/2)} \hat{p}_{C-1}(t, s, \mathbf{x}, \mathbf{y}), \end{aligned}$$

which is integrable only if for all $j \in \llbracket l, n \rrbracket$, $-(l-1/2)-1/2+\beta_j(j-1/2) > -1 \iff \beta_j > ((2l-2)/(2j-1))$. This condition actually holds if for any $i \in \llbracket 1, n \rrbracket$, $\beta_i > ((2i-2)/(2i-1))$ which is exactly the infimum assumed in (\mathbf{T}_β) . As we can see, there is no hope to obtain better thresholds with such a strategy. This is the reason why we said that these thresholds are *almost sharp* for the approach used here.

1.4.2 Back to the perturbative analysis

Let us now briefly explain what happens when one wants to control the approximation error in the forward parametrix expansion. Coming back to our general setting and denoting by \tilde{p}^m the transition density of our proxy, we obtain from the first order parametrix expansion the following representation for each regularized component \mathbf{u}_i^m , $i \in \llbracket 1, n \rrbracket$ of our solution \mathbf{U}^m of the systems (1.7): for all $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$

$$\mathbf{u}_i^m(t, \mathbf{x}) = \int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} \left\{ -\mathbf{F}_i^m(s, \mathbf{y}) + (L_s^m - \tilde{L}_s^m) \mathbf{u}_i^m(s, \mathbf{y}) \right\} \tilde{p}^m(t, s, \mathbf{x}, \mathbf{y}). \quad (1.21)$$

Above there is an additional term in the right hand side, in comparison with (1.14), which is precisely the approximation error due to the parametrix expansion. It thus appears that the solution has an implicit representation which makes its derivatives themselves appear. Hence, when differentiating the above representation to derive the estimate (1.9) in Theorem 2, we obtain bounds that depend themselves on the derivatives of the solution. We then have to estimate each derivative appearing in the right hand side and use a *circular argument*. Namely, when differentiating $\mathbf{u}_i^m(t, \mathbf{x})$, we will obtain the required estimate provided the multiplicative constants associated with the terms $\|\mathbf{D}\mathbf{u}_i^m\|_\infty$ and $\|D_1 \mathbf{D}\mathbf{u}_i^m\|_\infty$, that will appear in the corresponding upper-bound for the above right hand side, are small enough (see also Section 2 of [CdR17] and Section 2.2 below for details).

Moreover, as we have already seen, in order to smoothen the time singularity appearing when we apply a cross differentiation operator in the l^{th} and 1^{st} direction to the term $\int_t^T ds \int_{\mathbb{R}^{nd}} d\mathbf{y} (L_s^m - \tilde{L}_s^m) \mathbf{u}_i^m(s, \mathbf{y}) \tilde{p}^m(t, s, \mathbf{x}, \mathbf{y})$ corresponding to the approximation error, we will have to *center this term around the derivatives of the solution* itself (in the sense given in the above discussion). This procedure allows us, thanks to Taylor expansions, to weaken the singularities and provides integrable (in time) terms. The dramatic point is that, when doing so, our bound involves the cross derivatives $D_\ell \mathbf{D} \mathbf{u}_i^m$, $\ell \in \llbracket 1, n \rrbracket$ whose control in supremum norm is, as suggested by the discussion done in the explicit case of a simple degenerate Ornstein-Uhlenbeck process, definitely out of reach as soon as $\ell > 1$. In fact, as told by the results in [CdR17], the only thing we could hope is that the gradient in the degenerate directions viewed as a function of the degenerate variables, i.e. $\mathbf{D}_{2:n} \mathbf{u}_i^m(t, \mathbf{x}_1, \cdot) := (D_2 \mathbf{u}_i^m(t, \mathbf{x}_1, \cdot), \dots, D_n \mathbf{u}_i^m(t, \mathbf{x}_1, \cdot))$ for any $(t, \mathbf{x}_1) \in [0, T] \times \mathbb{R}^d$, belongs to an appropriate anisotropic Hölder space⁷ with regularity indexes strictly less than 1. Such spaces can as well be viewed as particular cases of anisotropic Besov spaces with corresponding positive regularity indexes. Thus, the generalized derivative of $\mathbf{D}_{2:n} \mathbf{u}_i^m(t, \mathbf{x}_1, \cdot)$ should belong to some anisotropic Besov space of negative regularity indexes, strictly bigger than -1 .

Here is the main novelty of our approach: to tackle this problem, our main idea, in order to balance the lack of differentiation property of the full gradient, consists in putting precisely in duality the anisotropic Besov norm with negative exponent of $(D_\ell \mathbf{D}_{2:n} \mathbf{u}_i^m(t, \mathbf{x}_1, \cdot))_{\ell \in \llbracket 2, n \rrbracket}$, with the corresponding anisotropic Besov norm with positive exponent of the remaining terms coming from the differentiation of (1.21), which in particular involve the coefficients of the operator $L^m - \tilde{L}^m$ and contain the time singularities coming from the derivatives of the frozen Gaussian kernel \tilde{p}^m . To the best of our knowledge, this approach is quite new in this parametrix setting and it appears to be very robust. We refer to the proof of the main Theorem 2 in Section 2 for details and to Proposition 3.6 in the book of Lemarié-Rieusset [LR02] for duality results on Besov spaces.

We are thus led to control on the one hand the Besov norm with negative exponent of the cross derivatives of the solution, see Lemma 10, and on the other hand the Besov norm with positive exponent of the remaining terms in (1.21) (involving the coefficients of the operator $L^m - \tilde{L}^m$), see Lemma 9. The first control (Besov norm with negative exponent) is crucial and appears to be quite delicate. Indeed, due to the implicit representation (1.21), this estimate also involves supremum norms of the full gradient $\mathbf{D} \mathbf{u}_i^m$ and of the cross derivatives themselves. This again reflects the *circular* nature of the arguments needed to derive the result.

To conclude this discussion on Besov duality, let us mention that a similar strategy has been implemented in our companion work [CdRHM18] in order to derive sharp Schauder estimates for the PDE (1.6) (with possibly non trivial final condition). Therein, since we were interested in controlling the Hölder norm of the solution, the duality was anyhow used the other way round: positive regularity indexes for the solution and negative ones for the remaining terms of the perturbative expansion.

Let us close this discussion coming back to Remark 1. As we emphasized, in comparison with the non-degenerate result, Theorem 1 should hold assuming that the drift \mathbf{F}_1 belongs to a suitable $L_q - L_p$ space w.r.t. time and the non-degenerate variable \mathbf{x}_1 . We are convinced that this is the case but we deliberately decide not to tackle this setting in order to keep this work shorter and more coherent. Indeed, in this case, the difficulty comes from the estimate on the second order derivative in the non-degenerate direction of the first component of the solution \mathbf{U}^m , namely $D_{\mathbf{x}_1} D_{\mathbf{x}_1} \mathbf{u}_1^m$ (which is a part of the main estimate (1.9) in Theorem 2). The point is to establish for this quantity an $L_q - L_p$ control. This cannot be derived from the previously described approach and requires harmonic analysis techniques (see also [KR05]). The main problem to establish the estimate is mainly due to the source term, which is actually \mathbf{F}_1 . To prove it, the main idea consists in exploiting the results of Menozzi [Men18] (where such an estimate is proved under the assumption that the drift is Lipschitz) through the tools developed in [CdRM17] (*backward* parametrix approach for drift \mathbf{F} whose first component may be in $L_q - L_p$ and the other ones in Hölder spaces). Then, the Zvonkin Transform should also be tuned a little bit following the strategy developed by Veretennikov (see e.g. [Ver80] and [FFPV17]). Such a program would surely toughen our paper without adding any surprising result and we prefer to focus on the novelty of the approach based on duality results for Besov spaces and the generalization of the strong uniqueness result to the whole chain (*i.e.* to any arbitrary $n \geq 1$) rather than drowning the reader into additional technical considerations.

2 Perturbation techniques for the PDE : proof of Theorem 2

In order to keep the notations as clear as possible, we forget the superscript m standing for the mollifying procedure and we suppose that the following assumptions hold:

⁷In other words, the regularity index depends on the considered variable

Assumption (AM). We say that assumption (AM) holds if the assumptions gathered in (A) hold true and the coefficients \mathbf{F} , a are infinitely differentiable functions with bounded derivatives of all order for a and greater than 1 for the coefficient \mathbf{F} .

In the whole section, we consider a fixed final time $T > 0$ which is meant to be *small*, i.e. $T \ll 1$. Let us consider for this section a generic PDE with generator corresponding to (1.5) and scalar source f having the same Hölder regularity than the drift terms in (1.1) (i.e. the scalar function f below can be any of the entries of the \mathbb{R}^d -valued $(\mathbf{F}_i)_{i \in [1, n]}$ in the dynamics (1.1)). Namely, we concentrate on

$$\begin{cases} (\partial_t u + L_t u)(t, \mathbf{x}) = -f(t, \mathbf{x}), & (t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}, \\ u(T, \mathbf{x}) = 0, \end{cases} \quad (2.1)$$

where $(L_t)_{t \geq 0}$ is defined in (1.5) and stands for the generator associated with (1.1) when the coefficients are smooth.

The key result to prove strong uniqueness for the SDE (1.1) is actually the following theorem from which one easily derive Theorem 2 for each component of the solution of the systems (1.6) and then the result in full generality.

Theorem 3 (Pointwise bounds for the derivatives of the PDE (2.1)). *There exists $\gamma := \gamma((\mathbf{A})) > 0$ and $C := C((\mathbf{A})) > 0$ s.t.*

$$\|\mathbf{D}u\|_\infty + \|\mathbf{D}(D_1 u)\|_\infty \leq CT^\gamma, \quad (2.2)$$

with obvious extension of the definition in (1.4) to the current scalar case.

The proof of Theorem 3 is performed in Section 2.2 through the forward parametrix approach consisting in considering a suitable *proxy* semi-group around which the initial solution of (2.1) can be expanded. To this end we first investigate in Section 2.1 below the linearized Gaussian process deriving from the dynamics in (1.1) which will provide the suitable model for the parametrix.

2.1 Gaussian proxy and associated controls

2.1.1 Linearization of the dynamics

Fix some freezing points $(\tau, \boldsymbol{\xi}) \in [0, T] \times \mathbb{R}^{nd}$. For fixed initial conditions $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$, a natural linearization associated with the mollified version of (1.1) writes

$$\begin{aligned} d\tilde{\mathbf{X}}_v^{(\tau, \boldsymbol{\xi})} &= [\mathbf{F}(v, \boldsymbol{\theta}_{v, \tau}(\boldsymbol{\xi})) + D\mathbf{F}(v, \boldsymbol{\theta}_{v, \tau}(\boldsymbol{\xi}))(\tilde{\mathbf{X}}_v^{(\tau, \boldsymbol{\xi})} - \boldsymbol{\theta}_{v, \tau}(\boldsymbol{\xi}))]dv + B\sigma(v, \boldsymbol{\theta}_{v, \tau}(\boldsymbol{\xi}))dW_v, \\ \forall v \in [t, s], \tilde{\mathbf{X}}_t^{(\tau, \boldsymbol{\xi})} &= \mathbf{x}, \end{aligned} \quad (2.3)$$

where

$$\dot{\boldsymbol{\theta}}_{v, \tau}(\boldsymbol{\xi}) = \mathbf{F}(v, \boldsymbol{\theta}_{v, \tau}(\boldsymbol{\xi})), \quad v \in [0, T], \quad \boldsymbol{\theta}_{\tau, \tau}(\boldsymbol{\xi}) = \boldsymbol{\xi}, \quad (2.4)$$

and $D\mathbf{F}(v, \cdot)$ denotes the subdiagonal of the Jacobian matrix $\mathbf{D}\mathbf{F}(v, \cdot)$. Namely, for $\mathbf{z} \in \mathbb{R}^{nd}$:

$$D\mathbf{F}(v, \mathbf{z}) = \begin{pmatrix} \mathbf{0}_{d,d} & \cdots & \cdots & \cdots & \mathbf{0}_{d,d} \\ D_{\mathbf{z}_1} \mathbf{F}_2(v, \mathbf{z}) & \mathbf{0}_{d,d} & \cdots & \cdots & \mathbf{0}_{d,d} \\ \mathbf{0}_{d,d} & D_{\mathbf{z}_2} \mathbf{F}_3(v, \mathbf{z}^{2:n}) & \mathbf{0}_{d,d} & \mathbf{0}_{d,d} & \vdots \\ \vdots & \mathbf{0}_{d,d} & \ddots & \vdots & \\ \mathbf{0}_{d,d} & \cdots & \mathbf{0}_{d,d} & D_{\mathbf{z}_{n-1}} \mathbf{F}_n(v, \mathbf{z}_{n-1}, \mathbf{z}_n) & \mathbf{0}_{d,d} \end{pmatrix}.$$

From the non-degeneracy of σ and Hörmander like condition, the Gaussian process defined by (2.5) admits a density $\tilde{p}^{(\tau, \boldsymbol{\xi})}(t, s, \mathbf{x}, \cdot)$ which is suitably controlled (see Proposition 5 below and for instance [DM10], [CdRM17]).

We explicitly integrate (2.3) to obtain for all $v \in [t, s]$:

$$\begin{aligned} \tilde{\mathbf{X}}_v^{(\tau, \boldsymbol{\xi})} &= \tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(v, t)\mathbf{x} + \int_t^v \tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(v, u) \left(\mathbf{F}(u, \boldsymbol{\theta}_{u, \tau}(\boldsymbol{\xi})) - D\mathbf{F}(u, \boldsymbol{\theta}_{u, \tau}(\boldsymbol{\xi}))\boldsymbol{\theta}_{u, \tau}(\boldsymbol{\xi}) \right) du \\ &\quad + \int_t^v \tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(v, u) B\sigma(u, \boldsymbol{\theta}_{u, \tau}(\boldsymbol{\xi})) dW_u \\ &=: \mathbf{m}_{v, t}^{(\tau, \boldsymbol{\xi})}(\mathbf{x}) + \int_t^v \tilde{\mathbf{R}}^{(\tau, \boldsymbol{\xi})}(v, u) B\sigma(u, \boldsymbol{\theta}_{u, \tau}(\boldsymbol{\xi})) dW_u, \end{aligned}$$

where $(\tilde{\mathbf{R}}^{(\tau, \xi)}(v, u))_{t \leq u, v \leq s}$ stands for the resolvent associated with the collection of partial gradients in $(D\mathbf{F}(v, \boldsymbol{\theta}_{v, \tau}(\boldsymbol{\xi})))_{v \in [t, s]}$ which satisfies for $v \in [t, s]$:

$$\partial_v \tilde{\mathbf{R}}^{(\tau, \xi)}(v, t) = D\mathbf{F}(v, \boldsymbol{\theta}_{v, \tau}(\boldsymbol{\xi})) \tilde{\mathbf{R}}^{(\tau, \xi)}(v, t), \quad \tilde{\mathbf{R}}^{(\tau, \xi)}(t, t) = \mathbf{I}_{nd \times nd}. \quad (2.5)$$

Note in particular that since the partial gradients are subdiagonal $\det(\tilde{\mathbf{R}}^{(\tau, \xi)}(v, t)) = 1$.

Also, for $v \in [t, s]$, we recall that $\mathbf{m}_{v, t}^{(\tau, \xi)}(\mathbf{x})$ stands for the mean of $\tilde{\mathbf{X}}_v^{(\tau, \xi)}$ and corresponds as well to the solution of (2.3) when $\sigma = 0$ when the starting point is \mathbf{x} . We write:

$$\tilde{\mathbf{X}}_v^{(\tau, \xi)} = \mathbf{m}_{v, t}^{(\tau, \xi)}(\mathbf{x}) + \int_t^v \tilde{\mathbf{R}}^{(\tau, \xi)}(v, u) B \sigma(u, \boldsymbol{\theta}_{u, \tau}(\boldsymbol{\xi})) dW_u, \quad v \in [t, s]. \quad (2.6)$$

Importantly, we point out that $\mathbf{x} \in \mathbb{R}^{nd} \mapsto \mathbf{m}_{v, t}^{(\tau, \xi)}(\mathbf{x})$ is **affine** w.r.t. the starting point \mathbf{x} . Precisely, for $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^{nd}$:

$$\mathbf{m}_{v, t}^{(\tau, \xi)}(\mathbf{x} + \mathbf{x}') = \tilde{\mathbf{R}}^{(\tau, \xi)}(v, t) \mathbf{x}' + \mathbf{m}_{v, t}^{(\tau, \xi)}(\mathbf{x}). \quad (2.7)$$

We first give in the next proposition a key estimates on the covariance matrix associated with (2.6) and its properties w.r.t. a suitable scaling of the system.

Proposition 4 (Good Scaling Properties of the Covariance Matrix). *The covariance matrix of $\tilde{\mathbf{X}}_v^{(\tau, \xi)}$ in (2.6) writes:*

$$\tilde{\mathbf{K}}_{v, t}^{(\tau, \xi)} := \int_t^v \tilde{\mathbf{R}}^{(\tau, \xi)}(v, u) B a(u, \boldsymbol{\theta}_{u, \tau}(\boldsymbol{\xi})) B^* \tilde{\mathbf{R}}^{(\tau, \xi)}(v, u)^* du.$$

Uniformly in $(\tau, \xi) \in [0, T] \times \mathbb{R}^{nd}$ and $s \in [0, T]$, it satisfies a good scaling property in the sense of Definition 3.2 in [DM10] (see also Proposition 3.4 of that reference). That is, for all fixed $T > 0$, there exists $C_{2.8} := C_{2.8}(\mathbf{A}, T) \geq 1$ s.t. for all $0 \leq t < v \leq s \leq T$, for all $(\tau, \xi) \in [0, T] \times \mathbb{R}^{nd}$:

$$\forall \boldsymbol{\zeta} \in \mathbb{R}^{nd}, \quad C_{2.8}^{-1} (v - t)^{-1} |\mathbb{T}_{v-t} \boldsymbol{\zeta}|^2 \leq \langle \tilde{\mathbf{K}}_{v, t}^{(\tau, \xi)} \boldsymbol{\zeta}, \boldsymbol{\zeta} \rangle \leq C_{2.8} (v - t)^{-1} |\mathbb{T}_{v-t} \boldsymbol{\zeta}|^2, \quad (2.8)$$

where we again use the notation introduced in (1.18) for the scaling matrix \mathbb{T}_{v-t} .

The proof of the above proposition readily follows from Proposition 3.3 and Lemma 3.6 in [DM10]. We now state some important density bounds for the linearized model.

Proposition 5 (Density of the linearized dynamics). *Under (\mathbf{A}) , we have that, for all $s \in (t, T]$ the random variable $\tilde{\mathbf{X}}_s^{(\tau, \xi)}$ in (2.6) admits a Gaussian density $\tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \cdot)$ which writes for all $\mathbf{y} \in \mathbb{R}^{nd}$:*

$$\tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}) := \frac{1}{(2\pi)^{\frac{nd}{2}} \det(\tilde{\mathbf{K}}_{s, t}^{(\tau, \xi)})^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \left\langle (\tilde{\mathbf{K}}_{s, t}^{(\tau, \xi)})^{-1} (\mathbf{m}_{s, t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y}), \mathbf{m}_{s, t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y} \right\rangle\right), \quad (2.9)$$

with $\tilde{\mathbf{K}}_{s, t}^{(\tau, \xi)}$ as in Proposition 4. Also, there exists $C := C(\mathbf{A}, T) > 0$ s.t. for all $l \in \llbracket 1, n \rrbracket$, $r \in \{0, 1\}$, we have:

$$\begin{aligned} |D_{\mathbf{x}_i} D_{\mathbf{x}_1}^r \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y})| &\leq \frac{C}{(s-t)^{\frac{nd}{2} + (l-\frac{1}{2}) + \frac{r}{2}}} \exp\left(-C^{-1} (s-t) |\mathbb{T}_{s-t}^{-1} (\mathbf{m}_{s, t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y})|^2\right) \\ &=: \frac{C}{(s-t)^{(l-\frac{1}{2}) + \frac{r}{2}}} \hat{p}_{C^{-1}}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}). \end{aligned} \quad (2.10)$$

Proof. Expression (2.9) readily follows from (2.5). The control (2.10) is then a direct consequence of Proposition 4 for $\alpha = 0$. Differentiating w.r.t. \mathbf{x} recalling from (2.7) that $\mathbf{x} \mapsto \mathbf{m}_{s, t}^{(\tau, \xi)}(\mathbf{x})$ is affine yields:

$$D_{\mathbf{x}_j} \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}) = - \left[[\tilde{\mathbf{R}}^{(\tau, \xi)}(s, t)]^* (\tilde{\mathbf{K}}_{s, t}^{(\tau, \xi)})^{-1} (\mathbf{m}_{s, t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y}) \right]_j \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}). \quad (2.11)$$

The point is now to use scaling arguments. We can first rewrite

$$[\tilde{\mathbf{R}}^{(\tau, \xi)}(s, t)]^* (\tilde{\mathbf{K}}_{s, t}^{(\tau, \xi)})^{-1} = (s-t) [\tilde{\mathbf{R}}^{(\tau, \xi)}(s, t)]^* \mathbb{T}_{s-t}^{-1} (\widehat{\mathbf{K}}_1^{s, t})^{-1} \mathbb{T}_{s-t}^{-1}, \quad (2.12)$$

where $\widehat{\mathbf{K}}_1^{s, t}$ is the covariance matrix of the rescaled process $((s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \tilde{\mathbf{X}}_{t+v(s-t)}^{t, \mathbf{x}})_{v \in [0, 1]}$ at time 1. From the good-scaling property of Proposition 4, it is plain to derive that $\widehat{\mathbf{K}}_1^{s, t}$ is a non-degenerate bounded matrix, i.e.

there exists $\hat{C} \geq 1$ s.t. for all $\zeta \in \mathbb{R}^{nd}$, $\hat{C}^{-1}|\zeta|^2 \leq \langle \widehat{\mathbf{K}}_1^{s,t} \zeta, \zeta \rangle \leq \hat{C}|\zeta|^2$. A similar rescaling argument yields on the deterministic system (2.5) of the resolvent yields that $\widehat{\mathbf{R}}^{(\tau, \xi)}(s, t)$ can also be written as:

$$[\tilde{\mathbf{R}}^{(\tau, \xi)}(s, t)]^* = \mathbb{T}_{s-t}^{-1} \left[\widehat{\mathbf{R}}^{(\tau, \xi), s, t}(1, 0) \right]^* \mathbb{T}_{s-t}, \quad (2.13)$$

where again $\widehat{\mathbf{R}}^{(\tau, \xi), s, t}(1, 0)$ is the resolvent at time 1 of the rescaled system $\mathbb{T}_{s-t}[\tilde{\mathbf{R}}^{(\tau, \xi)}(t + v(s-t), t)]^* \mathbb{T}_{s-t}^{-1} = \left(\left[\widehat{\mathbf{R}}^{(\tau, \xi), s, t}(v, 0) \right]^* \right)_{v \in [0,1]}$ associated with (2.5). From the analysis performed in Lemma 5.1 in [HM16] (see also the proof of Proposition 3.7 in [DM10]) one derives that there exists \hat{C}_1 s.t. for all $\zeta \in \mathbb{R}^{nd}$, $|\left[\widehat{\mathbf{R}}^{(\tau, \xi), s, t}(1, 0) \right]^* \zeta| \leq \hat{C}_1 |\zeta|$. Equations (2.11), (2.12) and (2.13) therefore yield:

$$\begin{aligned} & |D_{\mathbf{x}_j} \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y})| \\ & \leq (s-t)^{-j+\frac{1}{2}} \left| \left(\left[\widehat{\mathbf{R}}^{(\tau, \xi), s, t}(1, 0) \right]^* \left(\widehat{\mathbf{K}}_1^{s,t} \right)^{-1} \left((s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} (\mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y}) \right) \right)_j \right| \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}) \\ & \leq C (s-t)^{-j+\frac{1}{2}} (s-t)^{\frac{1}{2}} |\mathbb{T}_{s-t}^{-1} (\mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y})| \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}). \end{aligned}$$

From the explicit expression (2.9), Proposition 4 and the above equation, we eventually derive:

$$\begin{aligned} |D_{\mathbf{x}_j} \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y})| & \leq \frac{C}{(s-t)^{j-\frac{1}{2}}} \left((s-t)^{\frac{1}{2}} |\mathbb{T}_{s-t}^{-1} (\mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y})| \right) \\ & \quad \times \frac{1}{(s-t)^{\frac{n^2 d}{2}}} \exp \left(-C^{-1} (s-t) |\mathbb{T}_{s-t}^{-1} (\mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}) - \mathbf{y})|^2 \right) \\ & \leq \frac{C}{(s-t)^{j-\frac{1}{2}}} \tilde{p}_{C^{-1}}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}), \end{aligned}$$

up to a modification of C , which gives the statement for one partial derivative. The controls on the higher order derivatives are obtained similarly (see e.g. the proof of Lemma 5.5 of [DM10] for the bounds on $D_{\mathbf{x}_1}^2 \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y})$). \square

Now, let us specify a useful control involving the previous Gaussian kernel which will be exploited in some cancellation techniques.

Proposition 6. *For all $k \in \llbracket 1, n \rrbracket$, $0 \leq t \leq s \leq T$, $(\mathbf{x}, \xi) \in \mathbb{R}^{nd} \times \mathbb{R}^{nd}$, and $\mathbf{M} \in \mathbb{R}^d$ the following identity hold:*

$$\int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}) \langle \mathbf{M}, (\mathbf{y} - \mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}))_k \rangle d\mathbf{y} = \langle \mathbf{M}, \mathbf{1}_d \rangle, \quad \mathbf{1}_d = (1, \dots, 1) \in \mathbb{R}^d. \quad (2.14)$$

Proof. From Proposition 5, we have $\int_{\mathbb{R}^{nd}} \tilde{p}^{(\tau, \xi)}(t, s, \mathbf{x}, \mathbf{y}) (\mathbf{y} - \mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x}))_k d\mathbf{y} = \mathbf{0}_d$. Differentiating this expression w.r.t. \mathbf{x}_k and using the Leibniz formula (recalling as well the identity (2.7) which yields $D_{\mathbf{x}_k} [\mathbf{m}_{s,t}^{(\tau, \xi)}(\mathbf{x})]_k = (\widehat{\mathbf{R}}^{(\tau, \xi)}(s, t))_{k,k} = \mathbf{I}_{d,d}$) gives (2.14). \square

2.1.2 Associated inhomogeneous semi-group

Fix $t \in [0, T]$, $\xi \in \mathbb{R}^{nd}$. With the notations of the previous paragraph, we introduce the following inhomogeneous semi-group associated with (2.3) for $\tau = t$. Namely, for all $s \in (t, T]$, $g \in B_{\text{lin}}(\mathbb{R}^{nd}, \mathbb{R})$ (space of measurable functions with linear growth), $\mathbf{x} \in \mathbb{R}^{nd}$:

$$\tilde{P}_{s,t}^{\xi} g(\mathbf{x}) := \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{(t, \xi)}(t, s, \mathbf{x}, \mathbf{y}) g(\mathbf{y}). \quad (2.15)$$

From now on, we will write with a slight abuse of notation $\tilde{p}^{\xi}(t, s, \mathbf{x}, \mathbf{y}) := \tilde{p}^{(t, \xi)}(t, s, \mathbf{x}, \mathbf{y})$, i.e. we omit the freezing parameter in time when it corresponds to the considered starting time. One can derive from Proposition 5 the following important regularization result.

Lemma 7 (Regularization effects for the inhomogeneous semi-group). *Let $0 < \gamma \leq 1$. There exists $C := C((\mathbf{A}))$ such that for all $\ell \in \{0, 1\}$, $l \in \llbracket 1, n \rrbracket$, $k \in \llbracket 1, n \rrbracket$, $\mathbf{x} \in \mathbb{R}^{nd}$:*

$$|D_{\mathbf{x}_1}^{\ell} D_{\mathbf{x}_l} \tilde{P}_{s,t}^{\xi} \left(\left| (\cdot - \mathbf{m}_{s,t}^{(t, \xi)}(\mathbf{x}))_k \right|^{\gamma} \right) (\mathbf{x})|_{\xi=\mathbf{x}} \leq C (s-t)^{-\frac{\ell}{2} - (l-\frac{1}{2}) + \gamma(k-\frac{1}{2})}. \quad (2.16)$$

Let f be a ϑ -Hölder continuous functions where $\vartheta := (\vartheta_1, \dots, \vartheta_n) \in (0, 1]^n$ is a multi-index and for $i \in \llbracket 1, n \rrbracket$, ϑ_i stands for the Hölder regularity of f in the variable \mathbf{x}_i . The following result then holds.

- There exists $C := C(\mathbf{A})$ s.t. for all $l \in \{0, 1\}$, $l \in \llbracket 1, n \rrbracket$, $\mathbf{x} \in \mathbb{R}^{nd}$:

$$|D_{\mathbf{x}_1}^l D_{\mathbf{x}_l} \tilde{P}_{s,t}^{\xi} (f(\cdot) - f(\mathbf{m}_{s,t}^{(t,\xi)}(\mathbf{x}))) (\mathbf{x})|_{\xi=\mathbf{x}} \leq C \sum_{j=1}^n [f_j(s, \cdot)]_{\vartheta_j} (s-t)^{-\frac{\ell}{2} - (l - \frac{1}{2}) + \vartheta_j(j - \frac{1}{2})}, \quad (2.17)$$

where $[f_j(s, \cdot)]_{\vartheta_j} := \sup_{z \neq z' \in \mathbb{R}^d, (\mathbf{z}_{1:j-1}, \mathbf{z}_{j+1:n}) \in \mathbb{R}^{(n-1)d}} \frac{|f(\mathbf{z}_{1:j-1}, z', \mathbf{z}_{j+1:n}) - f(\mathbf{z}_{1:j-1}, z, \mathbf{z}_{j+1:n})|}{|z - z'|^{\vartheta_j}}$ stands for the Hölder continuity modulus of order ϑ_j of f in its j^{th} variable.

- Centering arguments. For all $l \in \llbracket 1, n \rrbracket$, $k \leq l$, it holds that:

$$D_{\mathbf{x}_l} \tilde{P}_{s,t}^{\xi} (f(\cdot)) (\mathbf{x}) = D_{\mathbf{x}_l} \tilde{P}_{s,t}^{\xi} (f(\cdot) - f(\cdot_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\xi))) (\mathbf{x}). \quad (2.18)$$

- As particular cases of the previous items, we have that there exists $C := C(\mathbf{A})$ s.t. for all $l \in \llbracket 1, n \rrbracket^2$, $\mathbf{x} \in \mathbb{R}^{nd}$:

$$\begin{aligned} |D_{\mathbf{x}_1} D_{\mathbf{x}_l} \tilde{P}_{s,t}^{\xi} f(\mathbf{x})|_{\xi=\mathbf{x}} &\leq C \sum_{j=l}^n [f_j(s, \cdot)]_{\vartheta_j} (s-t)^{-l + \vartheta_j(j - \frac{1}{2})}, \\ |D_{\mathbf{x}_l} \tilde{P}_{s,t}^{\xi} f(\mathbf{x})|_{\xi=\mathbf{x}} &\leq C \sum_{j=l}^n [f_j(s, \cdot)]_{\vartheta_j} (s-t)^{-(l - \frac{1}{2}) + \vartheta_j(j - \frac{1}{2})}. \end{aligned} \quad (2.19)$$

Proof. Let us first mention that identities (2.16) and (2.17) are direct consequences of Proposition 5.

Centering arguments like (2.18) will be a crucial tool in the analysis below. To justify such an identity, write:

$$\begin{aligned} D_{\mathbf{x}_l} \tilde{P}_{s,t}^{\xi} (f(\cdot_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\xi))) (\mathbf{x}) &= D_{\mathbf{x}_l} \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}^{\xi}(t, s, \mathbf{x}, \mathbf{y}) f(\mathbf{y}_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\xi)) \\ &= D_{\mathbf{x}_l} \int_{\mathbb{R}^{nd}} d\mathbf{y}_{1:l-1} f(\mathbf{y}_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\xi)) \int_{\mathbb{R}^{nd}} d\mathbf{y}_{l:n} \tilde{p}^{\xi}(t, s, \mathbf{x}, \mathbf{y}). \end{aligned}$$

Now, the structure of the linearized dynamics (2.3) yields that the variable \mathbf{x}_l only appears in $(\tilde{\mathbf{R}}^{(t,\xi)}(s, t)\mathbf{x})_{l:n}$ from its l^{th} to its n d -dimensional block. Therefore, setting e.g. $\tilde{\mathbf{y}}_{l:n} = \mathbf{y}_{l:n} - (\tilde{\mathbf{R}}^{(t,\xi)}(s, t)\mathbf{x})_{l:n}$, the integrated quantity $\int_{\mathbb{R}^{nd}} d\mathbf{y}_{l:n} \tilde{p}^{\xi}(t, s, \mathbf{x}, \mathbf{y})$ does not depend on \mathbf{x}_l anymore. This gives the statement (2.18).

Let us now prove (2.19). The idea is to use first a centering argument w.r.t. to the variables l to n . Namely,

$$\begin{aligned} & \left| D_{\mathbf{x}_1} D_{\mathbf{x}_l} \left[\tilde{P}_{s,t}^{\xi} f(s, \cdot) \right] (\mathbf{x}) \right| \\ &= \left| D_{\mathbf{x}_1} D_{\mathbf{x}_l} \left[\tilde{P}_{s,t}^{\xi} (f(s, \cdot) - f(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi))) \right] (\mathbf{x}) \right| \\ &\leq C (s-t)^{-(l-1/2)-1/2} \left[\tilde{P}_{s,t}^{\xi} \left(\sum_{j=l}^n [f_j(s, \cdot)]_{\vartheta_j} |(\cdot - \boldsymbol{\theta}_{s,t}(\xi))_j|^{\vartheta_j} \right) \right] (\mathbf{x}), \end{aligned}$$

where for a function $g \in B_{\text{lin}}(\mathbb{R}^d, \mathbb{R})$ we denote,

$$\hat{P}_{s,t}^{\xi} g(\mathbf{x}) = \int_{\mathbb{R}^{nd}} d\mathbf{y} \hat{p}_{C^{-1}}^{(t,\xi)}(t, s, \mathbf{x}, \mathbf{y}) g(\mathbf{y}) := \int_{\mathbb{R}^{nd}} d\mathbf{y} \hat{p}_{C^{-1}}^{\xi}(t, s, \mathbf{x}, \mathbf{y}) g(\mathbf{y}),$$

where $\hat{p}_{C^{-1}}^{(t,\xi)}(t, s, \mathbf{x}, \cdot) := \hat{p}_{C^{-1}}^{\xi}(t, s, \mathbf{x}, \cdot)$ is the density appearing in Proposition 5⁸. That is, $\hat{P}_{s,t}^{\xi}$ can somehow be viewed as the *pseudo* semi-group associated with the density $\hat{p}_{C^{-1}}^{(t,\xi)}$ appearing in Proposition 5. The dependence in C in $\hat{P}_{s,t}^{\xi}$ is then omitted for notational simplicity.

⁸we again use from now on the same notational simplification as for $\tilde{p}^{(t,\xi)}(t, s, \mathbf{x}, \cdot) := \tilde{p}^{\xi}(t, s, \mathbf{x}, \cdot)$, i.e. we omit the time superscript when it corresponds to the first argument of the density.

In the current regularized setting, it is plain to observe that taking $\boldsymbol{\xi} = \mathbf{x}$, $\mathbf{m}_{s,t}^{(t,\mathbf{x})}(\mathbf{x}) = \boldsymbol{\theta}_{s,t}(\mathbf{x})$. Therefore, from (2.10):

$$\begin{aligned}
\left| D_{\mathbf{x}_1} D_{\mathbf{x}_l} \left[\tilde{P}_{s,t}^{\boldsymbol{\xi}} f(s, \cdot) \right] (\mathbf{x}) \right|_{\boldsymbol{\xi}=\mathbf{x}} &\leq C \sum_{j=l}^n \int_{\mathbb{R}^{nd}} [f_j(s, \cdot)]_{\vartheta_j} d\mathbf{y} \frac{|(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\mathbf{x}))_j|^{\vartheta_j}}{(s-t)^l} \hat{p}_{C^{-1}}^{\mathbf{x}}(t, s, \mathbf{x}, \mathbf{y}) \\
&\leq C \sum_{j=l}^n [f_j(s, \cdot)]_{\vartheta_j} \int_{\mathbb{R}^{nd}} d\mathbf{y} \frac{|(\mathbf{y} - \mathbf{m}_{s,t}^{(t,\mathbf{x})}(\mathbf{x}))_j|^{\vartheta_j}}{(s-t)^l} \hat{p}_{C^{-1}}^{\mathbf{x}}(t, s, \mathbf{x}, \mathbf{y}) \\
&\leq C \sum_{j=l}^n [f_j(s, \cdot)]_{\vartheta_j} (s-t)^{-l+\vartheta_j(j-\frac{1}{2})}.
\end{aligned}$$

The control for $|D_{\mathbf{x}_l} \tilde{P}_{s,t}^{\boldsymbol{\xi}} f(\mathbf{x})|_{\boldsymbol{\xi}=\mathbf{x}}$ is derived similarly. \square

We state in the lemma below a useful control to obtain through Lemma 7 some smoothing effects for the degenerate part of the operator. The statement readily follow from (\mathbf{T}_β) .

Lemma 8. *From the smoothness assumption on the drift coefficient in (\mathbf{T}_β) , there exists $C := C((\mathbf{A}))$ s.t. for all $\ell \in \llbracket 2, n \rrbracket$, $k \geq \ell$, and for all $(s, \mathbf{x}, \boldsymbol{\xi}) \in [0, T] \times (\mathbb{R}^{nd})^2$:*

$$\begin{aligned}
&\left| \left(\mathbf{F}_\ell(s, \mathbf{y}_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})) - \mathbf{F}_\ell(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{\ell-1} \mathbf{F}_\ell(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{\ell-1} \right) \right| \\
&\leq C \left\{ \sum_{j=\ell}^{k-1} \left\{ [(\mathbf{F}_\ell)_j(s, \cdot)]_{\beta_j} |\mathbf{y} - \boldsymbol{\theta}_{s,t}(\mathbf{x})|_{\beta_j} \right\} + [(D_{\ell-1} \mathbf{F}_\ell)_{\ell-1}(s, \cdot)]_{\eta} |\mathbf{y} - \boldsymbol{\theta}_{s,t}(\mathbf{x})_{\ell-1}|^{1+\eta} \right\}.
\end{aligned}$$

2.2 Control of the sensitivities: proof of Theorem 3

To prove Theorem 3, the idea is to expand the solution of the PDE with regularized coefficients around a suitable proxy, as explained in Section 1.4. The proxy used here is the Gaussian process introduced in Section 2.1 for a suitable freezing parameter $\boldsymbol{\xi}$ to be specified later on. Then, the Duhamel formula (or first order parametrix expansion) yields:

$$\begin{aligned}
u(t, \mathbf{x}) &= \int_t^T ds \left[\tilde{P}_{s,t}^{\boldsymbol{\xi}} f(s, \cdot) \right] (\mathbf{x}) + \int_t^T ds \left[\tilde{P}_{s,t}^{\boldsymbol{\xi}} \left((L_s - \tilde{L}_s^{\boldsymbol{\xi}}) u \right) (s, \cdot) \right] (\mathbf{x}) \\
&= \int_t^T ds \left[\tilde{P}_{s,t}^{\boldsymbol{\xi}} f(s, \cdot) \right] (\mathbf{x}) + \int_t^T ds \left[\tilde{P}_{s,t}^{\boldsymbol{\xi}} \left(\langle \mathbf{F}_1(s, \cdot) - \mathbf{F}_1(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})), D_1 u(s, \cdot) \rangle \right) \right] (\mathbf{x}) \\
&\quad + \int_t^T ds \left[\tilde{P}_{s,t}^{\boldsymbol{\xi}} \left(\frac{1}{2} \text{Tr} \left[a(s, \cdot) - a(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) D_1^2 u(s, \cdot) \right] \right) \right] (\mathbf{x}) \\
&\quad + \int_t^T ds \left[\tilde{P}_{s,t}^{\boldsymbol{\xi}} \left(\sum_{i=2}^n \langle \mathbf{F}_i(s, \cdot) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) \right. \right. \\
&\quad \quad \left. \left. - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1}, D_i u(s, \cdot) \rangle \right) \right] (\mathbf{x}), \tag{2.20}
\end{aligned}$$

for any $\boldsymbol{\xi}$ in \mathbb{R}^{nd} .

To establish (2.2) we need to differentiate the above expression w.r.t. $(\mathbf{x}_l)_{l \in \llbracket 1, n \rrbracket}$ and then w.r.t. \mathbf{x}_1 and $(\mathbf{x}_l)_{l \in \llbracket 1, n \rrbracket}$ in order to obtain estimates depending only on known parameters in (\mathbf{A}) . Differentiating first this

expression w.r.t. $\mathbf{x}_l, l \in \llbracket 1, n \rrbracket$ we obtain:

$$\begin{aligned}
& D_{\mathbf{x}_l} u(t, \mathbf{x}) \\
&= \int_t^T ds D_{\mathbf{x}_l} \left[\tilde{P}_{s,t}^{\xi} f(s, \cdot) \right] (\mathbf{x}) + \int_t^T ds D_{\mathbf{x}_l} \left[\tilde{P}_{s,t}^{\xi} \left((L_s - \tilde{L}_s^{\xi}) u \right) (s, \cdot) \right] (\mathbf{x}) \\
&= \left\{ \int_t^T ds D_{\mathbf{x}_l} \left[\tilde{P}_{s,t}^{\xi} f(s, \cdot) \right] (\mathbf{x}) \right. \\
&\quad + \int_t^T ds D_{\mathbf{x}_l} \left[\tilde{P}_{s,t}^{\xi} \left(\langle (\mathbf{F}_1(s, \cdot) - \mathbf{F}_1(s, \boldsymbol{\theta}_{s,t}(\xi))), D_1 u(s, \cdot) \rangle \right) \right] (\mathbf{x}) \\
&\quad + \int_t^T ds D_{\mathbf{x}_l} \left[\tilde{P}_{s,t}^{\xi} \left(\frac{1}{2} \text{Tr} \left[\left(a(s, \cdot) - a(s, \boldsymbol{\theta}_{s,t}(\xi)) \right) D_1^2 u(s, \cdot) \right] \right) \right] (\mathbf{x}) \left. \right\} \\
&\quad + \left\{ \int_t^T ds D_{\mathbf{x}_l} \left[\tilde{P}_{s,t}^{\xi} \left(\sum_{i=2}^n \langle (\mathbf{F}_i(s, \cdot) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\xi)) \right. \right. \right. \\
&\quad \left. \left. \left. - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\xi)) (\cdot - \boldsymbol{\theta}_{s,t}(\xi))_{i-1} \right), D_i u(s, \cdot) \rangle \right) \right] (\mathbf{x}) \right\} \\
&=: \int_t^T ds H_l^{\xi}(s, \mathbf{x}) + \int_t^T ds I_l^{\xi}(s, \mathbf{x}). \tag{2.21}
\end{aligned}$$

The term $H_l^{\xi}(s, \mathbf{x})$ gathers all the derivatives of the solution w.r.t. the non-degenerate variables whereas $I_l^{\xi}(s, \mathbf{x})$ precisely gathers the derivatives w.r.t. the degenerate ones. We will now start from the representation (2.21) which we will again differentiate w.r.t. the non-degenerate variable \mathbf{x}_1 in order to prove the estimates of Theorem 3 concerning the second order derivatives which are the trickiest ones. Indeed, as it has been succinctly explained in Section 1.4, when differentiating the kernel associated with the frozen semi-group defined by (2.15) we generate an *a priori* not integrable time singularity which then needs to be smoothen by using, among others, tools developed in Lemma 7 (centering or cancellation arguments). The worst case then corresponds to the higher order of differentiation, namely $D_{\mathbf{x}_1} (D_{\mathbf{x}_1} u(t, \mathbf{x}))$ which, as suggested by Proposition 5, generates a time singularity of order $1/2 + (l - 1/2)$ in the time integrand of the r.h.s. of (2.21). We then only concentrate on this term and omit the proof of the statement concerning the boundedness of the gradient $D_{\mathbf{x}_l} u(t, \mathbf{x})$ which could be shown more directly.

The proof will be divided into two parts: we first handle the non-degenerate part of the operator (*i.e.* the estimate for $D_{\mathbf{x}_1} H_l(s, \mathbf{x})$) and then the degenerate part (*i.e.* the estimate for $D_{\mathbf{x}_1} I_l(s, \mathbf{x})$) which is a bit more involved.

Finally, before entering into the proof, we introduce some notations to ease the reading. We point out that the contributions H_l^{ξ}, I_l^{ξ} do depend on the freezing variable ξ . We omit this dependence for notational convenience when, as default ξ coincides with the spatial argument \mathbf{x} of the term. We write in this case $H_l^{\xi}(s, \mathbf{x})|_{\xi=\mathbf{x}} =: H_l(s, \mathbf{x})$ (resp. $I_l^{\xi}(s, \mathbf{x})|_{\xi=\mathbf{x}} =: I_l(s, \mathbf{x})$). The freezing parameter explicitly appears when a more careful choice is needed (see Section 3.2). Accordingly with the definition in (1.3), we will also use, with a slight abuse of notation recalling from (1.1) that the j^{th} variable appears with Hölder regularity β_j up to the level j of the chain, the notation

$$[(\mathbf{F}_{1:j})_j(s, \cdot)]_{\beta_j} := \max_{i \in \llbracket 1, j \rrbracket} [(\mathbf{F}_i)_j(s, \cdot)]_{\beta_j}. \tag{2.22}$$

Source term and non-degenerate part of the operator: estimates for $D_{\mathbf{x}_1} H_l(s, \mathbf{x})$. We first focus on the source term and the derivatives w.r.t. the non-degenerate variable \mathbf{x}_1 in (2.20) (three first terms in the r.h.s denoted from now on accordingly with (2.21) by $H_{l,1}(s, \mathbf{x}), H_{l,2}(s, \mathbf{x}), H_{l,3}(s, \mathbf{x})$).

For all $l \in \llbracket 1, n \rrbracket$, one readily derives from Lemma 7 (centering argument from the variables l to n) that for the source term:

$$|D_{\mathbf{x}_1} H_{l,1}(s, \mathbf{x})| := \left| D_{\mathbf{x}_1} D_{\mathbf{x}_l} \left[\tilde{P}_{s,t}^{\xi} f(s, \cdot) \right] (\mathbf{x}) \right|_{\xi=\mathbf{x}} \leq C \sum_{j=l}^n [f_j(s, \cdot)]_{\beta_j} (s-t)^{-l+\beta_j(j-1/2)}. \tag{2.23}$$

Those terms are integrable in time as soon as

$$\beta_j \left(j - \frac{1}{2} \right) - l > -1, \quad j \in \llbracket l, n \rrbracket \Leftrightarrow \beta_j \in \left(\frac{2j-2}{2j-1}; 1 \right]. \tag{2.24}$$

These thresholds are precisely those appearing in assumption (\mathbf{T}_β) and which lead to strong uniqueness for the associated SDE. Through our perturbative approach we will actually show that they are also precisely those leading to Theorem 3.

Similarly, from (2.20), (2.21), for the drift associated with the non-degenerate part, we first rewrite from the centering properties of Lemma 7:

$$\begin{aligned}
|D_{\mathbf{x}_1} H_{l,2}(s, \mathbf{x})| &:= \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[\tilde{P}_{s,t}^\xi \left(\langle \mathbf{F}_1(s, \cdot) - \mathbf{F}_1(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})), D_1 u(s, \cdot) \rangle \right) \right] (\mathbf{x}) \right| \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
&\leq \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[\tilde{P}_{s,t}^\xi \left(\langle \mathbf{F}_1(s, \cdot) - \mathbf{F}_1(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})), D_1 u(s, \cdot) \rangle \right) \right] (\mathbf{x}) \right| \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
&\quad + \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[\tilde{P}_{s,t}^\xi \left(\langle \mathbf{F}_1(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) - \mathbf{F}_1(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})), \right. \right. \right. \\
&\quad \left. \left. \left. (D_1 u(s, \cdot) - D_1 u(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi}))) \rangle \right) \right] (\mathbf{x}) \right| \Big|_{\boldsymbol{\xi}=\mathbf{x}}.
\end{aligned}$$

Thanks to Lemma 7, this again gives

$$\begin{aligned}
|D_{\mathbf{x}_1} H_{l,2}(s, \mathbf{x})| &\leq C \left(\sum_{j=l}^n \|D_1 u\|_\infty [(\mathbf{F}_1)_j(s, \cdot)]_{\beta_j} (s-t)^{-l+\beta_j(j-\frac{1}{2})} \right. \\
&\quad \left. + \sum_{j=l}^n \sum_{k=1}^{l-1} \|D_1 D_j u\|_\infty [(\mathbf{F}_1)_k(s, \cdot)]_{\beta_k} (s-t)^{-l+\beta_k(k-\frac{1}{2})+(j-\frac{1}{2})} \right), \tag{2.25}
\end{aligned}$$

leading precisely to the same integrability thresholds of equation (2.24) and assumption (\mathbf{T}_β) (as for the source term). The idea behind this control is crucial. We first handle, with the sole Hölder properties of drift and the supremum norm of $D_1 u$, the variables which are at a good smoothing scale w.r.t. the induced singularity. For the remaining term, which exhibits for the drift non-sufficient smoothing effects, we then additionally exploit a cancellation argument involving the gradient of the solution itself, which consequently makes the cross derivatives appear.

Eventually, we get for the diffusive part:

$$\begin{aligned}
&|D_{\mathbf{x}_1} H_{l,3}(s, \mathbf{x})| \\
&:= \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[\tilde{P}_{s,t}^\xi \left(\frac{1}{2} \text{Tr} \left[(a(s, \cdot) - a(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))) D_1^2 u(s, \cdot) \right] \right) \right] (\mathbf{x}) \right| \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
&= \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[\tilde{P}_{s,t}^\xi \left(\frac{1}{2} \text{Tr} \left[(a(s, \cdot) - a(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi}))) D_1^2 u(s, \cdot) \right] \right) \right] (\mathbf{x}) \right| \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
&\quad + \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[\tilde{P}_{s,t}^\xi \left(\frac{1}{2} \text{Tr} \left[(a(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) - a(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))) D_1^2 u(s, \cdot) \right] \right) \right] (\mathbf{x}) \right| \Big|_{\boldsymbol{\xi}=\mathbf{x}} \\
&=: |D_{\mathbf{x}_1} H_{l,31}(s, \mathbf{x})| + |D_{\mathbf{x}_1} H_{l,32}(s, \mathbf{x})|. \tag{2.26}
\end{aligned}$$

The term $D_{\mathbf{x}_1} H_{l,31}(s, \mathbf{x})$ is already centered at the *appropriate* scales, i.e. from variables l to n . We thus readily derive, similarly to the previous computations thanks to Lemma 7 and recalling that a is Lipschitz continuous, that:

$$|D_{\mathbf{x}_1} H_{l,31}(s, \mathbf{x})| \leq C \|D_1^2 u\|_\infty [a(s, \cdot)]_1 \sum_{j=l}^n (s-t)^{-l+(j-\frac{1}{2})} \leq C \|D_1^2 u\|_\infty [a(s, \cdot)]_1 (s-t)^{-\frac{1}{2}}, \tag{2.27}$$

which does not give a *critical* contribution w.r.t. the previously exhibited thresholds in (2.24) and (\mathbf{T}_β) . For the contribution $|D_{\mathbf{x}_1} H_{l,32}(s, \mathbf{x})|$ we use, in the same spirit as for $|D_{\mathbf{x}_1} H_{l,2}(s, \mathbf{x})|$, a centering argument and an

integration by parts to obtain:

$$\begin{aligned}
& |D_{\mathbf{x}_1} H_{l,32}(s, \mathbf{x})| \\
&= \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[\tilde{P}_{s,t}^{\xi} \left(\frac{1}{2} \text{Tr} \left[(a(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi)) - a(s, \boldsymbol{\theta}_{s,t}(\xi))) \right. \right. \right. \right. \\
&\quad \left. \left. \left. \times (D_1^2 u(s, \cdot) - D_1^2 u(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi))) \right] \right) \right] (\mathbf{x}) \right|_{\xi=\mathbf{x}} \\
&= \left| \frac{1}{2} D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left(\sum_{j=1}^d \int_{\mathbb{R}^{nd}} d\mathbf{y} \left\langle \partial_{\mathbf{y}_1^j} \left(\tilde{p}^{\xi}(t, s, \mathbf{x}, \mathbf{y}) (a_{j \cdot}(s, \mathbf{y}_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi)) - a_{j \cdot}(s, \boldsymbol{\theta}_{s,t}(\xi))) \right) \right. \right. \right. \\
&\quad \left. \left. \left. (D_1 u(s, \mathbf{y}) - D_1 u(s, \mathbf{y}_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi))) \right\rangle \right) \right|_{\xi=\mathbf{x}},
\end{aligned}$$

where $\mathbf{y}_1 = (\mathbf{y}_1^1, \dots, \mathbf{y}_1^d)$, $\partial_{\mathbf{y}_1^j}^j$ denotes the derivative w.r.t. to the j^{th} scalar entry of the non-degenerate variable \mathbf{y}_1 and $a_{j \cdot}$ denotes the j^{th} row of the diffusion matrix a . We therefore derive from Lemma 7 and the smoothness of a (by the Rademacher theorem, a is differentiable almost everywhere):

$$\begin{aligned}
|D_{\mathbf{x}_1} H_{l,32}(s, \mathbf{x})| &\leq C \| \mathbf{D} D_1 u(s, \cdot) \|_{\infty} \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(s-t)^l} \hat{p}_{C^{-1}}^{\mathbf{x}}(t, s, \mathbf{x}, \mathbf{y}) \\
&\quad \times \left(\frac{|\mathbf{y} - \boldsymbol{\theta}_{s,t}(\xi)|_{1:l-1}}{(s-t)^{\frac{1}{2}}} + 1 \right) |a(s, \cdot)_1| |\mathbf{y} - \boldsymbol{\theta}_{s,t}(\xi)|_{l:n}| \\
&\leq C \| \mathbf{D} D_1 u(s, \cdot) \|_{\infty} (s-t)^{-\frac{1}{2}}. \tag{2.28}
\end{aligned}$$

With the notations of (2.21), plugging (2.27), (2.28) into (2.26) and together with (2.25), (2.23), we eventually derive that there exists $\delta := \delta((\mathbf{A})) > 0$:

$$\left| \int_t^T ds D_{\mathbf{x}_1} H_l(s, \mathbf{x}) \right| \leq CT^{\delta} (\|D_1 u\|_{\infty} + \|\mathbf{D} D_1 u\|_{\infty}). \tag{2.29}$$

Degenerate part of the operator: estimates for $D_{\mathbf{x}_1} I_l(s, \mathbf{x})$. These are the most delicate terms to handle. Restarting from (2.21), we first write for all $l \in \llbracket 2, n \rrbracket$:

$$\begin{aligned}
& |D_{\mathbf{x}_1} I_l(s, \mathbf{x})| \\
&= \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[\tilde{P}_{s,t}^{\xi} \left(\sum_{i=2}^n \langle (\mathbf{F}_i(s, \cdot) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\xi)) - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\xi)) (\cdot - \boldsymbol{\theta}_{s,t}(\xi))_{i-1}), D_i u(s, \cdot) \rangle \right) \right] (\mathbf{x}) \right|_{\xi=\mathbf{x}},
\end{aligned}$$

which readily yields that

$$\begin{aligned}
& |D_{\mathbf{x}_1} I_l(s, \mathbf{x})| \\
&\leq \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[\tilde{P}_{s,t}^{\xi} \left(\sum_{i=l+1}^n \langle (\mathbf{F}_i(s, \cdot) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\xi)) - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\xi)) (\cdot - \boldsymbol{\theta}_{s,t}(\xi))_{i-1}), \right. \right. \right. \\
&\quad \left. \left. \left. D_i u(s, \cdot) \rangle \right) \right] (\mathbf{x}) \right|_{\xi=\mathbf{x}} \\
&\quad + \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[\tilde{P}_{s,t}^{\xi} \left(\sum_{i=2}^l \langle \mathbf{F}_i(s, \cdot) - \mathbf{F}_i(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi)), \right. \right. \right. \\
&\quad \left. \left. \left. D_i u(s, \cdot) \rangle \right) \right] (\mathbf{x}) \right|_{\xi=\mathbf{x}} \\
&\quad + \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[\tilde{P}_{s,t}^{\xi} \left(\sum_{i=2}^l \langle (\mathbf{F}_i(s, \cdot_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi)) - \mathbf{F}_i(\boldsymbol{\theta}_{s,t}(\xi)) \right. \right. \right. \\
&\quad \left. \left. \left. - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\xi)) (\cdot - \boldsymbol{\theta}_{s,t}(\xi))_{i-1}), D_i u(s, \cdot) \rangle \right) \right] (\mathbf{x}) \right|_{\xi=\mathbf{x}} \\
&=: |D_{\mathbf{x}_1} I_{1,l}(s, \mathbf{x})| + |D_{\mathbf{x}_1} I_{2,l}(s, \mathbf{x})| + |D_{\mathbf{x}_1} I_{3,l}(s, \mathbf{x})|. \tag{2.30}
\end{aligned}$$

We emphasize that the integrands $D_{\mathbf{x}_1} I_{1,l}(s, \mathbf{x})$ and $D_{\mathbf{x}_1} I_{2,l}(s, \mathbf{x})$ are already designed to smoothen the time singularities generated by the cross differentiation of the inhomogeneous semi-group w.r.t. the variables \mathbf{x}_l and \mathbf{x}_1 . Indeed, on the one hand

$$\begin{aligned} & |D_{\mathbf{x}_1} I_{1,l}(s, \mathbf{x})| \\ = & \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[\tilde{P}_{s,t}^{\xi} \left(\sum_{i=l+1}^n \langle (\mathbf{F}_i(s, \cdot) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\xi)) - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\xi))(\cdot - \boldsymbol{\theta}_{s,t}(\xi))_{i-1}), D_i u(s, \cdot) \rangle \right) \right] (\mathbf{x}) \right| \Big|_{\xi=\mathbf{x}}. \end{aligned}$$

Hence, Lemmas 7 and 8 yield:

$$\begin{aligned} & |D_{\mathbf{x}_1} I_{1,l}(s, \mathbf{x})| \tag{2.31} \\ \leq & C \|D_{l+1:n} u\|_{\infty} \sum_{j=l+1}^n \left\{ [(D_{j-1} \mathbf{F}_j)_{j-1}]_{\eta} (s-t)^{-l+(1+\eta)(j-3/2)} + [(\mathbf{F}_{l+1:n})_j]_{\beta_j} (s-t)^{-l+\beta_j(j-1/2)} \right\}, \end{aligned}$$

where $D_{l+1:n} u := (D_{l+1} u, \dots, D_n u)$ using as well the notation of (2.22) for the last inequality. Observe that the most singular terms in the previous bounds are those associated with the exponents $-l + \beta_j(j - 1/2)$. It turns out that they are indeed integrable under assumption (\mathbf{T}_{β}) giving again precisely the appropriate thresholds.

On the other hand, Lemma 7 also yields

$$\begin{aligned} |D_{\mathbf{x}_1} I_{2,l}(s, \mathbf{x})| &= \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[\tilde{P}_{s,t}^{\xi} \left(\sum_{i=2}^l \langle \mathbf{F}_i(s, \cdot) - \mathbf{F}_i(s, \cdot)_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi), D_i u(s, \cdot) \rangle \right) \right] (\mathbf{x}) \right| \Big|_{\xi=\mathbf{x}} \\ &\leq C \|D_{i:l-1} u\|_{\infty} \sum_{j=l}^n [(\mathbf{F}_{1:l})_j(s, \cdot)]_{\beta_j} (s-t)^{-l+\beta_j(j-1/2)}, \end{aligned} \tag{2.32}$$

and those terms are again integrable as soon as the thresholds of (\mathbf{T}_{β}) hold.

It hence remains to control the terms in $D_{\mathbf{x}_1} I_{3,l}$. These terms are the tricky ones since they are, *a priori*, not designed to smoothen the time singularities generated by the cross differentiation. Observe indeed that, if one tries to reproduce the above calculations, we obtain from Lemma 7 that

$$\begin{aligned} & |D_{\mathbf{x}_1} I_{3,l}(s, \mathbf{x})| \\ = & \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[\tilde{P}_{s,t}^{\xi} \left(\sum_{i=2}^l \langle (\mathbf{F}_i(s, \cdot)_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi)) - \mathbf{F}_i(\boldsymbol{\theta}_{s,t}(\xi)) \right. \right. \right. \\ & \quad \left. \left. \left. - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\xi))(\cdot - \boldsymbol{\theta}_{s,t}(\xi))_{i-1}), D_i u(s, \cdot) \rangle \right) \right] (\mathbf{x}) \right| \Big|_{\xi=\mathbf{x}} \\ \leq & C \|D_{2:l} u\|_{\infty} (s-t)^{-l} \\ & \times \left(\sum_{j=2}^l [(\mathbf{F}_{2:j})_j(s, \cdot)]_{\beta_j} (s-t)^{(j-\frac{1}{2})\beta_j} + [(D_{j-1} \mathbf{F}_j)_{j-1}(s, \cdot)]_{\eta} (s-t)^{(1+\eta)(j-\frac{3}{2})} \right), \\ \leq & C \|D_{2:l} u\|_{\infty} (s-t)^{-l} \left((s-t)^{\frac{3}{2}\beta_2} + (s-t)^{\frac{1}{2}(1+\eta)} \right), \end{aligned}$$

up to a modification of C and for T small enough. This leads, as soon as $l \geq 2$, to a time singularity which is not integrable. Indeed, $(1 + \eta)/2 < 1$ (recall that η is meant to be small). To overcome this problem, the idea consists in writing, thanks to Lemma 7,

$$\begin{aligned} & |D_{\mathbf{x}_1} I_{3,l}(s, \mathbf{x})| \\ = & \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[\tilde{P}_{s,t}^{\xi} \left(\sum_{i=2}^l \langle (\mathbf{F}_i(s, \cdot)_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi)) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\xi)) - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\xi))(\cdot - \boldsymbol{\theta}_{s,t}(\xi))_{i-1}, \right. \right. \right. \\ & \quad \left. \left. \left. (D_i u(s, \cdot) - D_i u(s, \cdot)_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\xi)) \rangle \right) \right] (\mathbf{x}) \right| \Big|_{\xi=\mathbf{x}} \end{aligned} \tag{2.33}$$

and to take advantage of the additional smoothing effect from the solution of the regularized PDE itself through the above contribution $D_i u(s, \cdot) - D_i u(s, \cdot)_{1:l-1, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})}$, $i \in \llbracket 2, l \rrbracket$. To do so we write, by expanding the gradients in (2.33) with the Taylor formula,

$$\begin{aligned}
& |D_{\mathbf{x}_1} I_{3,l}(s, \mathbf{x})| \\
&= \left| D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[\tilde{P}_{s,t}^{\boldsymbol{\xi}} \left(\sum_{i=2}^l \left\langle \left(\mathbf{F}_i(s, \cdot)_{1:l-1, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})} - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) \right. \right. \right. \right. \\
&\quad \left. \left. \left. - D_{i-1} \mathbf{F}_i(\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1} \right) \right. \right. \\
&\quad \left. \left. \left. \sum_{k=l}^n \int_0^1 d\lambda D_k D_i u(s, \cdot)_{1:l-1, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})} + \lambda(\cdot)_{l:n} - \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) (\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_k \right\rangle \right] (\mathbf{x}) \right| \\
&= \left| \sum_{i=2}^l \sum_{k=l}^n \int_0^1 d\lambda D_{\mathbf{x}_l} D_{\mathbf{x}_1} \left[\tilde{P}_{s,t}^{\boldsymbol{\xi}} \left(\left\langle \left(\mathbf{F}_i(s, \cdot)_{1:l-1, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})} - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) \right. \right. \right. \right. \right. \\
&\quad \left. \left. \left. - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1} \right) \right. \right. \\
&\quad \left. \left. \left. D_i D_k u(s, \cdot)_{1:l-1, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})} + \lambda(\cdot)_{l:n} - \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) (\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_k \right\rangle \right] (\mathbf{x}) \right|, \tag{2.34}
\end{aligned}$$

for any $l \in \llbracket 2, n \rrbracket$. Note that in the above expansion we used the Schwarz theorem to exchange the order of the derivatives. Indeed, we importantly point out that our approach *a priori* allows to control, in appropriate Besov spaces with negative exponents, the quantity $D_i D_k u$, only for $i \leq k$ (see Lemma 10 below). The above expansion hence allows to obtain the additional contribution $(\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_k$ which, thanks to Lemma 7, is precisely designed to smoothen the time singularity coming from the cross differentiation over the variables \mathbf{x}_l and \mathbf{x}_1 of the semigroup (notice indeed that $k \geq l$). Namely, introducing now for all $i \in \llbracket 2, l \rrbracket$, $k \in \llbracket l, n \rrbracket$, $(\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \in \mathbb{R}^{(n-1)d}$, $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{nd}$:

$$\begin{aligned}
& \Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})} : \mathbf{y}_i \in \mathbb{R}^d \mapsto \Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\mathbf{y}_i) = \\
& \left[D_{\mathbf{x}_l} D_{\mathbf{x}_1} \tilde{p}^{\boldsymbol{\xi}}(t, s, \mathbf{x}, \mathbf{y}) \otimes \left(\left(\mathbf{F}_i(\mathbf{y}_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) - \mathbf{F}_i(\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) \right. \right. \right. \\
& \quad \left. \left. \left. - D_{i-1} \mathbf{F}_i(\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1} \right) \right) ((\cdot - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_k)^* \right], \tag{2.35}
\end{aligned}$$

where the subscript $(1, l)$ in $\Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}$ is here to indicate the differentiation w.r.t. $D_{\mathbf{x}_l} D_{\mathbf{x}_1}$ acting on the frozen density (we will also use for the analysis the notation $\Psi_{i,l,k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}$ which indicates that the sole differentiation w.r.t. \mathbf{x}_l has to be taken into consideration, see Section 3.2 below). Pay attention that the above function is $(\mathbb{R}^d)^{\otimes 4}$ -valued.

We can expect a time smoothing effect for $\Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}$ of order $(s-t)^{-1/2+(i-1/2)\beta_i} + (s-t)^{-1/2+(i-1)-1/2(1+\eta)}$, where the exponent $-1/2$ derives from the differentiation $D_{\mathbf{x}_1}$. This fact can be easily deduced from Lemma 7 and the previous computations when $\boldsymbol{\xi} = \mathbf{x}$. With these notations at hand, it thus remains to control

$$\begin{aligned}
& |D_{\mathbf{x}_1} I_{3,l}(s, \mathbf{x})| \\
&\leq \sum_{i=2}^l \sum_{k=l}^n \left| \int_0^1 d\lambda \int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \right. \\
&\quad \left. \times \int_{\mathbb{R}^d} d\mathbf{y}_i \left\{ \left(\Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\mathbf{y}_i) : D_{\mathbf{y}_i} D_{\mathbf{y}_k} u(s, \mathbf{y}_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi}) + \lambda(\mathbf{y}_{l:n} - \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi}))) \right) \right\} \right|, \tag{2.36}
\end{aligned}$$

where “:” stands for the double tensor contraction, in terms of known parameters in (\mathbf{A}) . Unfortunately, as suggested by the computations in (2.23), the control in supremum norm of $(D_i D_k u)_{1 \leq i \leq k \leq n}$ is definitely out of reach. Recall indeed that in this equation (which corresponds to the case $i = 1$) the thresholds are sharp in order to retrieve an integrable singularity (see (2.24)).

Roughly speaking, what we can expect is that for fixed $(\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \in \mathbb{R}^{(n-1)d}$, the partial application $\mathbf{y}_i \mapsto D_k u^i(s, \mathbf{y}_i) := D_{\mathbf{y}_k} u(s, \mathbf{y})$, $\mathbf{y} := (\mathbf{y}_{1:i-1}, \mathbf{y}_i, \mathbf{y}_{i+1:n})$ is as smooth as the Green kernel involving the source

f in the indicated equation, i.e. $\tilde{G}^\xi f(s, \mathbf{y}) := \int_s^T dv \tilde{P}_{v,s}^\xi f(v, \mathbf{y})$. The specific Hölder modulus α_i^k of $\mathbf{y}_i \mapsto D_{\mathbf{y}_k} \tilde{G}^{i,\xi} f(s, \mathbf{y}_i) := D_{\mathbf{y}_k} \tilde{G}^\xi f(s, \mathbf{y})$ can be derived from Lemma 7 and we obtain that any α_i^k satisfying $\alpha_i^k < \frac{1-(1-\beta_k)(k-\frac{1}{2})}{i-\frac{1}{2}}$ is attainable. Our guess is hence that the same property should hold for $\mathbf{y}_i \mapsto D_{\mathbf{y}_k} u^i(s, \mathbf{y}_i)$.

Keeping in mind this objective, the next step consists in exploiting a Besov duality result on the variable \mathbf{y}_i recalling that $C_b^{\alpha_i^k}(\mathbb{R}^d, \mathbb{R}) = B_{\infty,\infty}^{\alpha_i^k}(\mathbb{R}^d, \mathbb{R})$, where from now on the notation $B_{p,q}^s$ stands for a Besov space with associated indexes p, q, s (see Triebel [Tri83] and Appendix C below). The indexes p, q denote the integrability parameters and s the smoothness one. However, instead of focusing on the Hölder modulus, we will concentrate on $\|D_i D_k u^i(s, \cdot)\|_{B_{\infty,\infty}^{\alpha_i^k-1}}$ (space of the derivatives of Hölder functions). This specific choice allows to naturally deal with the cancellation techniques which require to change the freezing points depending on the current regime of the underlying frozen heat-kernel (see the proof of Lemma 10 in Section 3.2 below). We also carefully mention that the quantity $\|D_i D_k u^i(s, \cdot)\|_{B_{\infty,\infty}^{\alpha_i^k-1}}$ is equivalent to $\|D_k u^i(s, \cdot)\|_{B_{\infty,\infty}^{\alpha_i^k}}$ (see Proposition 14).

A classical fact is that $B_{\infty,\infty}^{\alpha_i^k-1}$ and $B_{1,1}^{1-\alpha_i^k}$ can be put in duality, see e.g. Proposition 3.6 in Lemarié-Rieusset ([LR02]). Indeed, with the notations therein $B_{\infty,\infty}^{\alpha_i^k-1}$ is the dual of $B_{1,1}^{1-\alpha_i^k} = \tilde{B}_{1,1}^{1-\alpha_i^k}$ where $\tilde{B}_{1,1}^{1-\alpha_i^k}$ denotes the closure of the Schwartz class \mathcal{S} in $B_{1,1}^{1-\alpha_i^k}$ (see also Theorem 4.1.3 in Adams Hedberg [AH96] for the density of \mathcal{S} in $B_{1,1}^{1-\alpha_i^k}$). Exploiting this fact, we then derive from (2.36) and the multi-linearity of the tensors involved that:

$$\begin{aligned}
& |D_{\mathbf{x}_1} I_{3,l}(s, \mathbf{x})| \\
& \leq C \sum_{i=2}^l \sum_{k=l}^n \int_0^1 d\lambda \int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \\
& \quad \times \left\{ \left\| \left[\Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\cdot) \right] \right\|_{B_{1,1}^{1-\alpha_i^k}} \left\| \mathbf{y}_i \mapsto D_i D_k u(s, \mathbf{y}_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi}) + \lambda(\mathbf{y}_{l:n} - \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi}))) \right\|_{B_{\infty,\infty}^{\alpha_i^k-1}} \right\} \\
& \leq C \sum_{i=2}^l \sum_{k=l}^n \int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \left\{ \left\| \left[\Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\cdot) \right] \right\|_{B_{1,1}^{1-\alpha_i^k}} \right. \\
& \quad \times \left[\sup_{\mathbf{z}_j, j \in [1, n], j \neq i} \left\| D_i D_k u(s, \mathbf{z}_{1:i-1}, \cdot, \mathbf{z}_{i+1:n}) \right\|_{B_{\infty,\infty}^{\alpha_i^k-1}} \mathbb{I}_{i \leq l-1} \right. \\
& \quad \left. \left. + \sup_{\mathbf{z}_j, j \in [1, n], j \neq i, \lambda \in [0, 1]} \left\| D_i D_k u(s, \mathbf{z}_{1:i-1}, \boldsymbol{\theta}_{s,t}^i(\boldsymbol{\xi}) + \lambda(\cdot - \boldsymbol{\theta}_{s,t}^i(\boldsymbol{\xi})), \mathbf{z}_{i+1:n}) \right\|_{B_{\infty,\infty}^{\alpha_i^k-1}} \mathbb{I}_{i=l} \right] \right\}. \tag{2.37}
\end{aligned}$$

Observe now that, for any $\lambda \in [0, 1]$ and the norm equivalence of Proposition 14, we have:

$$\begin{aligned}
& \left\| D_i D_k u(s, \mathbf{z}_{1:i-1}, \boldsymbol{\theta}_{s,t}^i(\boldsymbol{\xi}) + \lambda(\cdot - \boldsymbol{\theta}_{s,t}^i(\boldsymbol{\xi})), \mathbf{z}_{i+1:n}) \right\|_{B_{\infty,\infty}^{\alpha_i^k-1}} \\
& \leq C \left\| D_k u(s, \mathbf{z}_{1:i-1}, \boldsymbol{\theta}_{s,t}^i(\boldsymbol{\xi}) + \lambda(\cdot - \boldsymbol{\theta}_{s,t}^i(\boldsymbol{\xi})), \mathbf{z}_{i+1:n}) \right\|_{B_{\infty,\infty}^{\alpha_i^k}} \\
& \leq C \left\| D_k u(s, \mathbf{z}_{1:i-1}, \cdot, \mathbf{z}_{i+1:n}) \right\|_{B_{\infty,\infty}^{\alpha_i^k}},
\end{aligned}$$

exploiting the scaling and shift invariance properties of the Hölder modulus. Using again the norm equivalence of Proposition 14, we eventually derive:

$$\begin{aligned}
& \left\| D_i D_k u(s, \mathbf{z}_{1:i-1}, \boldsymbol{\theta}_{s,t}^i(\boldsymbol{\xi}) + \lambda(\cdot - \boldsymbol{\theta}_{s,t}^i(\boldsymbol{\xi})), \mathbf{z}_{i+1:n}) \right\|_{B_{\infty,\infty}^{\alpha_i^k-1}} \\
& \leq C^2 \left\| D_i D_k u(s, \mathbf{z}_{1:i-1}, \cdot, \mathbf{z}_{i+1:n}) \right\|_{B_{\infty,\infty}^{\alpha_i^k-1}}. \tag{2.38}
\end{aligned}$$

In other words, the dilations-translations of the variable \mathbf{y}_i , that appear when $i = l$ in (2.37), do not affect the

regularity estimate. We therefore obtain:

$$\begin{aligned}
& |D_{\mathbf{x}_1} I_{3,l}(s, \mathbf{x})| \\
& \leq C \sum_{i=2}^l \sum_{k=l}^n \int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}) \left\{ \left\| \left[\Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\cdot) \right] \right\|_{B_{1,1}^{1-\alpha_i^k}} \right. \\
& \quad \left. \times \sup_{\mathbf{z}_j, j \in \llbracket 1, n \rrbracket, j \neq i} \left\| D_i D_k u(s, \mathbf{z}_{1:i-1}, \cdot, \mathbf{z}_{i+1:n}) \right\|_{B_{\infty, \infty}^{\alpha_i^k - 1}} \right\}.
\end{aligned} \tag{2.39}$$

To conclude this proof we now need the following results whose proofs are postponed to the next subsection:

Lemma 9. *Let $l \in \llbracket 2, n \rrbracket$, $i \in \llbracket 2, l \rrbracket$ and $k \in \llbracket l, n \rrbracket$ and let $\Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ be the function defined by (2.35). There exist $C := C(\mathbf{A}) > 0$, $\alpha_i^k := \frac{1+\frac{\eta}{4}}{2i-1} < \frac{1-(1-\beta_k)(k-\frac{1}{2})}{i-\frac{1}{2}}$, $\gamma_i^k = \gamma_i := \frac{1}{2} + \eta(i - \frac{3}{2}) > \frac{1}{2}$ such that*

$$\left\| \left[\Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\cdot) \right] \right\|_{B_{1,1}^{1-\alpha_i^k}} \leq C(s-t)^{-\frac{3}{2} + \gamma_i^k} \hat{q}_{c \setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})),$$

where, with the notations of Proposition 5,

$$\hat{q}_{c \setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) := \int_{\mathbb{R}^d} \hat{p}_c^{\mathbf{x}}(t, s, \mathbf{x}, \mathbf{y}) d\mathbf{y}_i = \prod_{j \in \llbracket 1, n \rrbracket, j \neq i}^n \mathcal{N}_{c(s-t)^{2j-1}}((\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})_j), \tag{2.40}$$

denoting as well for $a > 0$, $z \in \mathbb{R}^d$, by $\mathcal{N}_a(z) = \frac{1}{(2\pi a)^{\frac{d}{2}}} \exp\left(-\frac{|z|^2}{2a}\right)$ the standard Gaussian density of \mathbb{R}^d with covariance matrix $a\mathbf{I}_d$. In the above control $c = C^{-1}$.

Lemma 10. *Let u be the solution of (2.1).*

There exists $C := C(\mathbf{A}) > 0$ such that for all $i \leq k \in \llbracket 2, n \rrbracket^2$ and $\alpha_i^k := \frac{1+\frac{\eta}{4}}{2i-1} < \frac{1-(1-\beta_k)(k-\frac{1}{2})}{i-\frac{1}{2}}$,

$$\sup_{\mathbf{y}_j, j \in \llbracket 1, n \rrbracket, j \neq i} \left\| D_i D_k u(s, \mathbf{y}_{1:i-1}, \cdot, \mathbf{y}_{i+1:n}) \right\|_{B_{\infty, \infty}^{\alpha_i^k - 1}} \leq C(\|\mathbf{D}u\|_{\infty} + \|\mathbf{D}D_1 u\|_{\infty}). \tag{2.41}$$

Remark 3 (About Lemmas 9 and 10). *There are some specific points to be emphasized about the indicated lemmas: We point out that the parameters γ_i^k , α_i^k actually do not depend on the index k . We keep the notations to remember that they are associated with the Hölder regularity of $D_{\mathbf{x}_k} u$ w.r.t. to its i^{th} variable which we choose to investigate in the corresponding negative Besov space for technical reasons. We also emphasize that the norm*

$\left\| D_i D_k u(s, \mathbf{y}_{1:i-1}, \cdot, \mathbf{y}_{i+1:n}) \right\|_{B_{\infty, \infty}^{\alpha_i^k - 1}}$ is actually equivalent $\left\| D_k u(s, \mathbf{y}_{1:i-1}, \cdot, \mathbf{y}_{i+1:n}) \right\|_{B_{\infty, \infty}^{\alpha_i^k}}$ which corresponds to

the usual Hölder norm on $C_b^{\alpha_i^k}$ (see Proposition 14 below). This can be seen again as a specific feature appearing when dealing with strong uniqueness through the Zvonkin approach. The regularity is associated with a variable independently of the level in the chain (1.1).

Remark 4 (Hölder moduli of the gradients). *Once Lemma 10 is proved, it follows from the Schwarz theorem that for all $i \leq k$, $D_k D_i u = D_i D_k u$ is also controlled in $B_{\infty, \infty}^{\alpha_i^k - 1}$. From the norm equivalence of Proposition 14 this means that $D_k u$ is Hölder continuous in its i^{th} variable for all i in $\llbracket 1, n \rrbracket$.*

We can then deduce, using Lemma 9 and Lemma 10 that

$$|D_{\mathbf{x}_1} I_{3,l}(s, \mathbf{x})| \leq C \sum_{i=2}^l \sum_{k=l}^n (s-t)^{-\frac{3}{2} + \gamma_i^k} (\|\mathbf{D}u\|_{\infty} + \|\mathbf{D}D_1 u\|_{\infty}), \tag{2.42}$$

which are integrable terms since $\gamma_i^k > \frac{1}{2}$. With the notations of (2.21), (2.30), we eventually derive from (2.42), (2.32), (2.31) that there exists $\gamma := \gamma(\mathbf{A}) > 0$ such that:

$$\left| \int_t^T ds D_{\mathbf{x}_1} I_l(s, \mathbf{x}) \right| \leq CT^{\gamma} (\|\mathbf{D}u\|_{\infty} + \|\mathbf{D}D_1 u\|_{\infty}). \tag{2.43}$$

Conclusion. Bringing together (2.43) and (2.29) yields for all $l \in \llbracket 1, n \rrbracket$ and all $(t, \mathbf{x}) \in [0, T]$:

$$|D_{\mathbf{x}_l} D_{\mathbf{x}_1} u(t, \mathbf{x})| \leq C(T^\gamma + T^\delta) \left(\|\mathbf{D}u\|_\infty + \|\mathbf{D}D_1 u\|_\infty \right). \quad (2.44)$$

It is clear that the previous analysis can be reproduced without differentiating w.r.t. \mathbf{x}_1 , leading to improved singularity exponents (see also the proof of Lemma 10 which somehow exactly explicit these computations). We therefore get:

$$|D_{\mathbf{x}_l} u(t, \mathbf{x})| \leq C(T^{\gamma'} + T^{\delta'}) \left(\|\mathbf{D}u\|_\infty + \|\mathbf{D}D_1 u\|_\infty \right), \quad (2.45)$$

for some positive exponents γ', δ' (with $\gamma' > \gamma$, $\delta' > \delta$).

Taking the time-space supremum in the l.h.s of (2.44) and (2.45), recalling as well that T is meant to be small, i.e. s.t. $4CT^{\delta \wedge \gamma} \leq \frac{1}{2}$, we derive:

$$\|\mathbf{D}u\|_\infty + \|\mathbf{D}D_1 u\|_\infty \leq 2C(T^\gamma + T^\delta).$$

This concludes the proof. \square

3 Estimates in Besov norm

This section is dedicated to the proofs of the main technical results needed to obtain Theorem 3. Namely, we prove the Besov estimates of Lemmas 9 and 10.

3.1 Proof of Lemma 9

We will here exploit the thermic characterization of Besov spaces (see Chapter 2.6.4 in [Tri83]) which is also recalled in Appendix C.

From (C.1), we are thus led to estimate, for any $l \in \llbracket 2, n \rrbracket$, $i \in \llbracket 2, l-1 \rrbracket$ and $k \in \llbracket l, n \rrbracket$:

$$\int_0^1 \frac{dv}{v} v^{1 + \frac{\alpha_i^k - 1}{2}} \|\partial_v h_v \star \Psi_{i, (l, 1), k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\cdot)\|_{L^1(\mathbb{R}^d, \mathbb{R})}.$$

We split the time integral in the above equation into two parts writing:

$$\begin{aligned} & \int_0^{(s-t)^{\rho_{i,k}}} dv v^{\frac{\alpha_i^k - 1}{2}} \|\partial_v h_v \star \Psi_{i, (l, 1), k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\cdot)\|_{L^1(\mathbb{R}^d, \mathbb{R})} \\ & + \int_{(s-t)^{\rho_{i,k}}}^1 dv v^{\frac{\alpha_i^k - 1}{2}} \|\partial_v h_v \star \Psi_{i, (l, 1), k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\cdot)\|_{L^1(\mathbb{R}^d, \mathbb{R})} \\ & =: \mathbf{Lower} \left[\Psi_{i, (l, 1), k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})} \right] + \mathbf{Upper} \left[\Psi_{i, (l, 1), k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})} \right], \end{aligned} \quad (3.1)$$

for a parameter $\rho_{i,k} > 0$ to be specified. The term **Upper** corresponding to the upper-part of the integral w.r.t. v does not involve singularities. We will use this fact to calibrate the associated parameter $\rho_{i,k}$ in order to match the integrability constraint

$$\mathbf{Upper} \left[\Psi_{i, (l, 1), k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})} \right] \leq \frac{C}{(s-t)^{1 + \frac{1}{2} - \gamma_i^k}} \hat{q}_{c \setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})), \quad (3.2)$$

where $\hat{q}_{c \setminus i}$ has been defined in (2.40) and $\gamma_i^k > 1/2$ in order to obtain a time integrable singularity. For this term, we will only use crude upper-bounds on the derivatives of the heat-kernel and the coefficients satisfying (\mathbf{T}_β) . On the other hand, the contribution **Lower** in (3.1) precisely contains the singularities w.r.t. v . It is therefore crucial to use there suitable cancellation tools. The point will then be to prove that the associated estimates are compatible with the upper-bound in equation (3.2).

We now write:

$$\begin{aligned} & \mathbf{Upper} \left[\Psi_{i, (l, 1), k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})} \right] \\ & = \int_{(s-t)^{\rho_{i,k}}}^1 dv v^{\frac{\alpha_i^k - 1}{2}} \|\partial_v h_v \star \Psi_{i, (l, 1), k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\cdot)\|_{L^1(\mathbb{R}^d, \mathbb{R})} \\ & = \int_{(s-t)^{\rho_{i,k}}}^1 dv v^{\frac{\alpha_i^k - 1}{2}} \int_{\mathbb{R}^d} dz \left| \int_{\mathbb{R}^d} \partial_v h_v(z - \mathbf{y}_i) \Psi_{i, (l, 1), k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\mathbf{y}_i) d\mathbf{y}_i \right|. \end{aligned}$$

Recall from the definition in (2.35) that $\Psi_{i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(\mathbf{y}_i)$ is $(\mathbb{R}^d)^{\otimes 4}$ -valued. To proceed with the computations we assume w.l.o.g. for the rest of the proof that $d = 1$ to avoid tensor notations for simplicity. Writing explicitly the function $\Psi_{i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(\mathbf{y}_i)$ leads to:

$$\begin{aligned} & \text{Upper} \left[\Psi_{i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})} \right] \\ & \leq \int_{(s-t)^{\rho_{i,k}}}^1 dv v^{\frac{\alpha_i^k - 1}{2}} \int_{\mathbb{R}^d} dz \left| \int_{\mathbb{R}^d} \partial_v h_v(z - \mathbf{y}_i) \left(D_{\mathbf{x}_1} D_{\mathbf{x}_i} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) \left[((\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}) - \mathbf{y})_k \right. \right. \right. \right. \\ & \quad \left. \left. \left. \left(\mathbf{F}_i(s, \mathbf{y}_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1} \right) \right] \right) \right|_{\boldsymbol{\xi}=\mathbf{x}}. \end{aligned}$$

From Lemma 8 and Proposition 5, we derive there exists a $C := C((\mathbf{A}), T) > 0$ such that introducing $\hat{q}_c(t, s, \mathbf{x}, \mathbf{y}) = \tilde{p}_c^\mathbf{x}(t, s, \mathbf{x}, \mathbf{y})$, $c = C^{-1}$:

$$\begin{aligned} & \text{Upper} \left[\Psi_{i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})} \right] \\ & \leq C \int_{(s-t)^{\rho_{i,k}}}^1 dv v^{\frac{\alpha_i^k - 1}{2}} \int_{\mathbb{R}^d} dz \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z - \mathbf{y}_i)}{v} \frac{\hat{q}_c(t, s, \mathbf{x}, \mathbf{y})}{(s-t)^{(l-\frac{1}{2})+\frac{1}{2}}} |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})_k| \\ & \quad \times \left\{ \sum_{j=i}^{l-1} \left\{ |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})_j|^{\beta_j} \right\} + |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \mathbf{y})_{i-1}|^{1+\eta} \right\} \\ & \leq C \int_{(s-t)^{\rho_{i,k}}}^1 dv v^{\frac{\alpha_i^k - 1}{2}} \int_{\mathbb{R}^d} dz \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z - \mathbf{y}_i)}{v} \frac{\hat{q}_c(t, s, \mathbf{x}, \mathbf{y})}{(s-t)^{\frac{1}{2}}} \\ & \quad \times \left\{ \sum_{j=i}^{l-1} (s-t)^{\beta_j(j-1/2)} + (s-t)^{(1+\eta)(i-3/2)} \right\} \\ & \leq C \hat{q}_{c \setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \int_{(s-t)^{\rho_{i,k}}}^1 dv v^{-\frac{3}{2} + \frac{\alpha_i^k}{2}} (s-t)^{-1/2} \\ & \quad \times \left\{ \sum_{j=i}^{l-1} (s-t)^{\beta_j(j-1/2)} + (s-t)^{(1+\eta)(i-3/2)} \right\} \\ & \leq C \hat{q}_{c \setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) (s-t)^{[-\frac{1}{2} + \frac{\alpha_i^k}{2}] \rho_{i,k} - \frac{1}{2}} \left((s-t)^{\beta_i(i-1/2)} + (s-t)^{(1+\eta)(i-3/2)} \right), \end{aligned}$$

recalling that the lower bound of $\beta_j(j-1/2)$ is increasing for the last inequality (recall indeed that we assumed that $\beta_j \in \left(\frac{2j-2}{2j-1}, 1 \right]$). We now want to choose the threshold $\rho_{i,k}$ in order to match the integrability condition in (3.2). This amounts to write:

$$-1 - \frac{1}{2} + \gamma_i^k = \left[-\frac{1}{2} + \frac{\alpha_i^k}{2} \right] \rho_{i,k} - \frac{1}{2} + \left(\beta_i(i-1/2) \wedge (1+\eta)(i-3/2) \right).$$

Since this condition should hold for any $\beta_i \in \left(\frac{2i-2}{2i-1}, 1 \right]$ and since the parameter $\eta \in (0, 1)$ is small (see (\mathbf{H}_η)) we have $\beta_i(i-1/2) \wedge (1+\eta)(i-3/2) = (1+\eta)(i-3/2)$. The above condition rewrites:

$$-1 - \frac{1}{2} + \gamma_i^k = \left[-\frac{1}{2} + \frac{\alpha_i^k}{2} \right] \rho_{i,k} - \frac{1}{2} + (1+\eta)(i-3/2).$$

Our global integrability constraint associated with the i^{th} variable in the k^{th} derivative writes:

$$-1 - \frac{1}{2} + \gamma_i^k = \left[-\frac{1}{2} + \frac{\alpha_i^k}{2} \right] \rho_{i,k} - \frac{1}{2} + (1+\eta)(i-3/2) > -1, \quad (3.3)$$

which gives

$$\rho_{i,k} < \frac{(1+\eta)(2i-3) + 1}{1 - \alpha_i^k}. \quad (3.4)$$

It therefore remains to check that such a choice is compatible with the time integral part for $v \in [0, (s-t)^{\rho_{i,k}}]$ in the thermic characterisation of the Besov norm. We point out that for this term it is absolutely essential to get rid of the exponent v^{-1} coming from the upper-bound of the thermic heat-kernel, i.e. $\partial_v h_v(z - \mathbf{y}_i)$. In order to get an integrable singularity in v , we need to decrease the crude upper-bound on $\partial_v h_v(z - \mathbf{y}_i)$. This is done through cancellation techniques exploiting the smoothness properties of $\Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}$.

To investigate $\mathbf{Lower} \left[\Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})} \right]$ let us first recall from the definition in (2.35) that for all $k \in \llbracket l, n \rrbracket$:

$$\begin{aligned} & \Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\mathbf{y}_i) \\ &= D_{\mathbf{x}_l} D_{\mathbf{x}_1} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) \left[(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_k \left(\mathbf{F}_i(\mathbf{y}_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) - \mathbf{F}_i(\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) \right) \right. \\ & \quad \left. - D_{i-1} \mathbf{F}_i(\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1} \right]. \end{aligned} \quad (3.5)$$

Let us now specify the dependence w.r.t. \mathbf{y}_i of the previous expression in function of the considered indexes $i \in \llbracket 1, l \rrbracket, l \in \llbracket 2, n \rrbracket, k \in \llbracket l, n \rrbracket$. This will be useful to develop corresponding adapted cancellation arguments.

Observe first that the dependence in \mathbf{y}_i appears in (3.5) for any $i \in \llbracket 1, l \rrbracket, l \in \llbracket 2, n \rrbracket, k \in \llbracket l, n \rrbracket$ through the term $D_{\mathbf{x}_1 \mathbf{x}_2} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y})$.

For the term into brackets we distinguish two cases. If $i \leq l-1$, and then $k > i$, the only bracket term containing \mathbf{y}_i is the one associated with \mathbf{F}_i . If now $i = l \leq k$, for the contribution $i = l = k$, the only term into brackets in (3.5) that will also depend on \mathbf{y}_i is $(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_i$ and the cancellation arguments need to be slightly modified. Eventually, when $i = l < k$, there will be no dependence on \mathbf{y}_i for the terms into brackets.

- First case $i < l \leq k$ ($i \in \llbracket 1, l-1 \rrbracket, l \in \llbracket 2, n \rrbracket, k \in \llbracket l, n \rrbracket$).

With the notations of (3.1), we write:

$$\begin{aligned} & \mathbf{Lower} \left[\Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})} \right] \\ &= \int_0^{(s-t)^{\rho_{i,k}}} dv v^{\frac{\alpha_k-1}{2}} \int_{\mathbb{R}^d} dz \left| \int_{\mathbb{R}^d} d\mathbf{y}_i \partial_v h_{cv}(z - \mathbf{y}_i) \right. \\ & \quad \left. (\Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\mathbf{y}_i) - \Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(z)) \right| \\ &=: \int_0^{(s-t)^{\rho_{i,k}}} dv v^{\frac{\alpha_k-1}{2}} \int_{\mathbb{R}^d} dz \left| (\mathcal{F}_{1,i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})} + \mathcal{F}_{2,i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})})(v, z) \right|, \end{aligned} \quad (3.6)$$

where:

$$\begin{aligned} & \mathcal{F}_{1,i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(v, z) \\ &:= \int_{\mathbb{R}^d} \partial_v h_v(z - \mathbf{y}_i) \left(D_{\mathbf{x}_1} D_{\mathbf{x}_l} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) \right. \\ & \quad \left. \left[(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_k \left(\mathbf{F}_i(s, \mathbf{y}_{1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) - \mathbf{F}_i(s, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) \right) \right] \right] d\mathbf{y}_i, \end{aligned} \quad (3.7)$$

with a slight abuse of notation when $i = l-1$ and

$$\begin{aligned} & \mathcal{F}_{2,i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(v, z) \\ &:= \int_{\mathbb{R}^d} \partial_v h_v(z - \mathbf{y}_i) \left[D_{\mathbf{x}_1} D_{\mathbf{x}_l} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_1} D_{\mathbf{x}_l} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}) \right] \left[(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_k \right. \\ & \quad \left. \times \left(\mathbf{F}_i(s, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:l-1}, \boldsymbol{\theta}_{s,t}^{l:n}(\boldsymbol{\xi})) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1} \right) \right] d\mathbf{y}_i. \end{aligned} \quad (3.8)$$

Write now from (3.8), Proposition 5 and Lemma 8 (recalling as well that we took $\boldsymbol{\xi} = \mathbf{x}$ for the current

analysis):

$$\begin{aligned}
& |\mathcal{F}_{2,i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(v,z)| \\
& \leq C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z-\mathbf{y}_i)}{v} \int_0^1 d\lambda \frac{\hat{q}_c(t,s,\mathbf{x},\mathbf{y}_{1:i-1},z+\lambda(\mathbf{y}_i-z),\mathbf{y}_{i+1:n})}{(s-t)^{(i-\frac{1}{2})+(l-\frac{1}{2})+\frac{1}{2}}} |(\mathbf{y}-\boldsymbol{\theta}_{s,t}(\mathbf{x}))_k| |\mathbf{y}_i-z| \\
& \quad \times \left(\left| \mathbf{F}_i(s,\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:l-1},\boldsymbol{\theta}_{s,t}^{l:n}(\mathbf{x})) - \mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\mathbf{x})) - D_{i-1}\mathbf{F}_i(s,\boldsymbol{\theta}_{s,t}(\mathbf{x}))(\mathbf{y}-\boldsymbol{\theta}_{s,t}(\mathbf{x}))_{i-1} \right| \right) \\
& \leq C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z-\mathbf{y}_i)}{v^{\frac{1}{2}}} \int_0^1 d\lambda \frac{\hat{q}_c(t,s,\mathbf{x},\mathbf{y}_{1:i-1},z+\lambda(\mathbf{y}_i-z),\mathbf{y}_{i+1:n})}{(s-t)^{(i-\frac{1}{2})+\frac{1}{2}}} \\
& \quad \times \left\{ |z-(\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i|^{\beta_i} + \sum_{j=i+1}^{l-1} (s-t)^{\beta_j(j-1/2)} + (s-t)^{(1+\eta)(i-3/2)} \right\},
\end{aligned}$$

where for the second inequality, we used that for $k \geq l > i$, $|(\mathbf{y}-\boldsymbol{\theta}_{s,t}(\mathbf{x}))_k|(s-t)^{-(l-1/2)} \leq |(\mathbf{y}-\boldsymbol{\theta}_{s,t}(\mathbf{x}))_k|(s-t)^{-(k-1/2)}$ which can be absorbed by the k^{th} variables of \hat{q}_c .

Writing now for any $\lambda \in [0, 1]$,

$$|z - \boldsymbol{\theta}_{s,t}(\mathbf{x})_i| \leq \lambda|z - \mathbf{y}_i| + |z + \lambda(\mathbf{y}_i - z) - (\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i|,$$

we thus derive

$$\begin{aligned}
& |\mathcal{F}_{2,i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(v,z)| \\
& \leq C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z-\mathbf{y}_i)}{v^{\frac{1}{2}}} \int_0^1 d\lambda \hat{q}_c(t,s,\mathbf{x},\mathbf{y}_{1:i-1},z+\lambda(\mathbf{y}_i-z),\mathbf{y}_{i+1:n}) \\
& \quad \times \left(v^{\frac{\beta_i}{2}} (s-t)^{-(i-\frac{1}{2})-\frac{1}{2}} + (s-t)^{-(i-\frac{1}{2})-\frac{1}{2}+\beta_i(i-1/2)} \right. \\
& \quad \left. + \sum_{j=i+1}^{l-1} (s-t)^{-(i-\frac{1}{2})-\frac{1}{2}+\beta_j(j-1/2)} + (s-t)^{-(i-\frac{1}{2})-\frac{1}{2}+(1+\eta)(i-3/2)} \right). \\
& \leq C \hat{q}_{c \setminus i}(t,s,\mathbf{x},\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}) \int_0^1 d\lambda \int_{\mathbb{R}^d} d\mathbf{y}_i h_{cv}(z-\mathbf{y}_i) \mathcal{N}_{c(s-t)^{2i-1}}(z+\lambda(\mathbf{y}_i-z) - (\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i) \\
& \quad \times v^{-\frac{1}{2}} \left(v^{\frac{\beta_i}{2}} (s-t)^{-(i-\frac{1}{2})-\frac{1}{2}} + (s-t)^{-(i-\frac{1}{2})-\frac{1}{2}+(1+\eta)(i-3/2)} \right), \tag{3.9}
\end{aligned}$$

recalling for the last inequality that for all j in $\llbracket i, l-1 \rrbracket$, $\beta_j(j-1/2) > j-1 > i-3/2$ and η is supposed to be a small parameter.

Consider now \mathcal{F}_1 ,

$$\begin{aligned}
|\mathcal{F}_{1,i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}(v,z)| & \leq C \int_{\mathbb{R}^d} \frac{h_{cv}(z-\mathbf{y}_i)}{v} \frac{\hat{q}_c(t,s,\mathbf{x},\mathbf{y})}{(s-t)^{(l-\frac{1}{2})+\frac{1}{2}}} |z-\mathbf{y}_i|^{\beta_i} |(\mathbf{y}-\boldsymbol{\theta}_{s,t}(\mathbf{x}))_k| d\mathbf{y}_i \\
& \leq C \int_{\mathbb{R}^d} \frac{h_{cv}(z-\mathbf{y}_i)}{v^{1-\frac{\beta_i}{2}}} \frac{\hat{q}_c(t,s,\mathbf{x},\mathbf{y})}{(s-t)^{\frac{1}{2}}} d\mathbf{y}_i. \tag{3.10}
\end{aligned}$$

From (3.9) and (3.10) we derive, with the notation introduced in (2.40):

$$\begin{aligned}
& \|\partial_v h_v \star \Psi_{i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})}\|_{L^1(\mathbb{R}^d, \mathbb{R})} \\
& \leq C \hat{q}_{c \setminus i}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})) \left\{ \frac{1}{v^{1-\frac{\beta_i}{2}}(s-t)^{\frac{1}{2}}} + \left(\frac{v^{-\frac{1+\beta_i}{2}}}{(s-t)^i} + \frac{v^{-\frac{1}{2}}}{(s-t)^{i-(1+\eta)(i-3/2)}} \right) \right. \\
& \quad \left. \times \int_0^1 d\lambda \int_{\mathbb{R}^d} dz \int_{\mathbb{R}^d} d\mathbf{y}_i h_{cv}(z-\mathbf{y}_i) \mathcal{N}_{c(s-t)^{2i-1}}(z+\lambda(\mathbf{y}_i-z) - (\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i) \right\} \\
& \leq C \hat{q}_{c \setminus i}(t,s,\mathbf{x},(\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n})) \left(\frac{1}{v^{1-\frac{\beta_i}{2}}(s-t)^{\frac{1}{2}}} + \frac{v^{-\frac{1+\beta_i}{2}}}{(s-t)^i} + \frac{v^{-\frac{1}{2}}}{(s-t)^{i-(1+\eta)(i-3/2)}} \right),
\end{aligned}$$

using the change of variable $(w_1, w_2) = (z - \mathbf{y}_i, z + \lambda(\mathbf{y}_i - z) - (\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i)$ for the last inequality.

For $\rho_{i,k}$ chosen as in (3.4) one gets from the definition in (3.1):

$$\begin{aligned} & \mathbf{Lower} \left[\Psi_{i,(l,1),k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})} \right] \\ & \leq C \hat{q}_{c \setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \int_0^{(s-t)^{\rho_{i,k}}} \frac{dv}{v} v^{\frac{\alpha_i^k}{2}} \\ & \quad \times \left(\frac{1}{v^{\frac{1-\beta_i}{2}} (s-t)^{\frac{1}{2}}} + \frac{v^{\frac{\beta_i}{2}}}{(s-t)^i} + \frac{1}{(s-t)^{i-(1+\eta)(i-3/2)}} \right) \\ & =: C \hat{q}_{c \setminus i}(t, s, \mathbf{x}, (\mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n})) \mathcal{B}_{1, \rho_{i,k}}(t, s). \end{aligned}$$

It therefore remains to prove that, if $\mathcal{B}_{1, \rho_{i,k}}(t, s) \leq (s-t)^{-1-1/2+\gamma_i^k}$ for $\gamma_i^k > 0$ then:

$$-1 - \frac{1}{2} + \gamma_i^k > -1 \iff \gamma_i^k > \frac{1}{2}. \quad (3.11)$$

Write:

$$\begin{aligned} \mathcal{B}_{1, \rho_{i,k}}(t, s) & \leq \int_0^{(s-t)^{\rho_{i,k}}} dv \left(\frac{v^{-\frac{3}{2} + \frac{\alpha_i^k + \beta_i}{2}}}{(s-t)^{\frac{1}{2}}} + \frac{v^{-1 + \frac{\alpha_i^k + \beta_i}{2}}}{(s-t)^i} + \frac{v^{-1 + \frac{\alpha_i^k}{2}}}{(s-t)^{\frac{3}{2} - \eta(i - \frac{3}{2})}} \right) \\ & \leq C \left((s-t)^{\rho_{i,k} \frac{\alpha_i^k + \beta_i - 1}{2} - \frac{1}{2}} + (s-t)^{\rho_{i,k} \frac{\alpha_i^k + \beta_i}{2} - i} + (s-t)^{\rho_{i,k} \frac{\alpha_i^k}{2} - \frac{3}{2} + \eta(i - \frac{3}{2})} \right). \end{aligned} \quad (3.12)$$

Let us check condition (3.11) is satisfied. Actually, the first two terms in (3.12) yield negligible contributions. Recall indeed from the statement of the lemma that the parameter α_i^k must be chosen so that $\alpha_i^k < \frac{1-(1-\beta_k)(k-\frac{1}{2})}{i-\frac{1}{2}} := \bar{\alpha}_i^k$. Since $\beta_k \in (\frac{2k-2}{2k-1}, 1]$, $\bar{\alpha}_i^k = \frac{2-(1-\beta_k)(2k-1)}{2i-1} > \frac{3-2k+(2k-2)}{2i-1} = \frac{1}{2i-1}$. Thus,

$$\alpha_i^k = \frac{1 + \frac{\eta}{4}}{2i-1}, \quad (3.13)$$

is an admissible choice (recall indeed that we chose $\eta < \inf_{j \in [2, n]} \{\beta_j - \frac{2j-2}{2j-1}\}$ in assumption (\mathbf{H}_η)). It gives in particular that $\alpha_i^k + \beta_i > \frac{1 + \frac{\eta}{4} + (2i-2)}{2i-1} = 1 + \frac{\eta}{2i-1}$. Hence,

$$\rho_{i,k} \frac{\alpha_i^k + \beta_i - 1}{2} - \frac{1}{2} > -\frac{1}{2}$$

which already provides a regularizing term in time for the first term in the r.h.s. of (3.12). Now from (3.4), $\rho_{i,k} < \frac{(1+\eta)(2i-3)+1}{1-\alpha_i^k} =: \bar{\rho}_{i,k}$. The previous choice for α_i^k gives

$$\bar{\rho}_{i,k} = \frac{2i-2 + \eta(2i-3)}{\frac{2i-1-(1+\frac{\eta}{4})}{2i-1}} = (2i-1) \frac{2i-2 + \eta(2i-3)}{2i-2 - \frac{\eta}{4}} > (2i-1),$$

and

$$\rho_{i,k} = 2i-1 \quad (3.14)$$

is an admissible choice. We therefore get for the exponent of the second term in (3.12):

$$\rho_{i,k} \frac{\alpha_i^k + \beta_i}{2} - i > \frac{1}{2} \left(\rho_{i,k} \left(1 + \frac{\eta}{4} \right) - 2i \right) > -\frac{1}{2}.$$

Eventually, for the exponent of the third contribution in (3.12), for the previous choice of $\rho_{i,k} = 2i-1$, we get

$$\rho_{i,k} \frac{\alpha_i^k}{2} - \frac{3}{2} + \eta \left(i - \frac{3}{2} \right) = \frac{1}{2} (2i-1) \frac{1 + \frac{\eta}{4}}{2i-1} - \frac{3}{2} + \eta \left(i - \frac{3}{2} \right) = -1 + \frac{\eta}{8} + \eta \left(i - \frac{3}{2} \right) > -1, \quad (3.15)$$

which means that criterion (3.11) is indeed satisfied, even though if this last contribution is rather critical in order to obtain the required smoothing effect with $\gamma_i^k = 1/2 + \eta(i-3/2)$. This concludes the proof of Lemma

9 for the indexes $i \in \llbracket 1, l-1 \rrbracket$.

• Second case $i = l \leq k$. Let us begin with the case $i = l < k$ for which the only contribution in \mathbf{y}_i in (3.5) appears through $D_{\mathbf{x}_1} D_{\mathbf{x}_i} \tilde{p}(t, s, \mathbf{x}, \mathbf{y})$. Hence, for such indexes,

$$\text{Lower} \left[\Psi_{i, (i,1), k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})} \right] = \int_0^{(s-t)^{\rho_{i,k}}} dv v^{\frac{\alpha_i^k - 1}{2}} \int_{\mathbb{R}^d} dz |\mathcal{F}_{2, i, (i,1), k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(v, z)|,$$

where:

$$\begin{aligned} & \mathcal{F}_{2, i, (i,1), k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(v, z) \\ = & \int_{\mathbb{R}^d} d\mathbf{y}_i \partial_v h_v(z - \mathbf{y}_i) \left[D_{\mathbf{x}_1} D_{\mathbf{x}_i} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_1} D_{\mathbf{x}_i} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}) \right] \left[(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_k \right. \\ & \left. \times \left(\mathbf{F}_i(s, \mathbf{y}_{1:i-1}, \boldsymbol{\theta}_{s,t}^{i:n}(\boldsymbol{\xi})) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1} \right) \right]. \end{aligned} \quad (3.16)$$

The previous analysis for this term can be reproduced adapting the computations leading to (3.9). Precisely,

$$\begin{aligned} & |\mathcal{F}_{2, i, (i,1), k}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(v, z)| \\ \leq & C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z - \mathbf{y}_i)}{v} \int_0^1 d\lambda \frac{\hat{q}_c(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z + \lambda(\mathbf{y}_i - z), \mathbf{y}_{i+1:n})}{(s-t)^{(i-\frac{1}{2}) + (l-\frac{1}{2}) + \frac{1}{2}}} |(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\mathbf{x}))_k| |\mathbf{y}_i - z| \\ & \times \left| (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\mathbf{x}))_{i-1} \right|^{1+\eta} \\ \leq & C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z - \mathbf{y}_i)}{v^{\frac{1}{2}}} \int_0^1 d\lambda \frac{\hat{q}_c(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z + \lambda(\mathbf{y}_i - z), \mathbf{y}_{i+1:n})}{(s-t)^{(i-\frac{1}{2}) + \frac{1}{2}}} (s-t)^{(1+\eta)(i-3/2)}. \end{aligned}$$

This contribution has already been analyzed and yields the expected integrable singularity in time.

Let us now focus on the remaining case $i = l = k$, for which we write:

$$\begin{aligned} & \text{Lower} \left[\Psi_{i, (i,1), i}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})} \right] \\ = & \int_0^{(s-t)^{\rho_{i,i}}} dv v^{\frac{\alpha_i^i - 1}{2}} \int_{\mathbb{R}^d} dz \left| \int_{\mathbb{R}^d} d\mathbf{y}_i \partial_v h_v(z - \mathbf{y}_i) \right. \\ & \left. \left(\Psi_{i, (i,1), i}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(\mathbf{y}_i) - \Psi_{i, (i,1), i}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(z) \right) \right| \\ = & \int_0^{(s-t)^{\rho_{i,i}}} dv v^{\frac{\alpha_i^i - 1}{2}} \int_{\mathbb{R}^d} dz \left| \int_{\mathbb{R}^d} d\mathbf{y}_i \partial_v h_v(z - \mathbf{y}_i) \right. \\ & \left[D_{\mathbf{x}_1} D_{\mathbf{x}_i} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y})(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_i - D_{\mathbf{x}_1} D_{\mathbf{x}_i} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n})(z - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_i \right] \\ & \left(\mathbf{F}_i(s, \mathbf{y}_{1:i-1}, \boldsymbol{\theta}_{s,t}^{i:n}(\boldsymbol{\xi})) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1} \right) \left. \right| \\ =: & \int_0^{(s-t)^{\rho_{i,i}}} dv v^{\frac{\alpha_i^i - 1}{2}} \int_{\mathbb{R}^d} dz |\mathcal{F}_{2, i, (i,1), i}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})} + \mathcal{F}_{3, i, (i,1), i}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(v, z)|, \end{aligned} \quad (3.17)$$

where $\mathcal{F}_{2, i, (i,1), i}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(v, z)$ is obtained from (3.16) taking $k = i$ and

$$\begin{aligned} & \mathcal{F}_{3, i, (i,1), i}^{(s, \mathbf{y}_{1:i-1}, \mathbf{y}_{i+1:n}), (t, \mathbf{x})}(v, z) \\ := & \int_{\mathbb{R}^d} d\mathbf{y}_i \partial_v h_v(z - \mathbf{y}_i) \left(D_{\mathbf{x}_1} D_{\mathbf{x}_i} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}_{1:i-1}, z, \mathbf{y}_{i+1:n}) \right. \\ & \left. \left[(\mathbf{y}_i - z) \left(\mathbf{F}_i(s, \mathbf{y}_{1:i-1}, \boldsymbol{\theta}_{s,t}^{i:n}(\boldsymbol{\xi})) - \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{i-1} \mathbf{F}_i(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))(\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1} \right) \right] \right). \end{aligned} \quad (3.18)$$

Write first:

$$\begin{aligned}
& |\mathcal{F}_{2,i,(i,1),i}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}),(\mathbf{t},\mathbf{x})}(v,z)| \\
& \leq C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z-\mathbf{y}_i)}{v} \int_0^1 d\lambda \frac{\hat{q}_c(t,s,\mathbf{x},\mathbf{y}_{1:i-1},\mathbf{y}_i+\lambda(z-\mathbf{y}_i),\mathbf{y}_{i+1:n})}{(s-t)^{(i-\frac{1}{2})+(i-\frac{1}{2})+\frac{1}{2}}} |(\mathbf{y}-\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i| |\mathbf{y}_i-z| \\
& \quad \times \left| (\mathbf{y}-\boldsymbol{\theta}_{s,t}(\mathbf{x}))_{i-1} \right|^{1+\eta} \\
& \leq C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z-\mathbf{y}_i)}{v^{\frac{1}{2}}} \int_0^1 d\lambda \hat{q}_c(t,s,\mathbf{x},\mathbf{y}_{1:i-1},\mathbf{y}_i+\lambda(\mathbf{y}_i-z),\mathbf{y}_{i+1:n}) \\
& \quad \times \left(\frac{|\mathbf{y}_i+\lambda(z-\mathbf{y}_i)-(\boldsymbol{\theta}_{s,t}(\mathbf{x}))_i|+\lambda|z-\mathbf{y}_i|}{(s-t)^{2i-\frac{1}{2}}} \right) (s-t)^{(i-\frac{3}{2})(1+\eta)}.
\end{aligned}$$

We thus derive:

$$\begin{aligned}
& |\mathcal{F}_{2,i,(i,1),i}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}),(\mathbf{t},\mathbf{x})}(v,z)| \\
& \leq C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z-\mathbf{y}_i)}{v^{\frac{1}{2}}} \int_0^1 d\lambda \hat{q}_c(t,s,\mathbf{x},\mathbf{y}_{1:i-1},\mathbf{y}_i+\lambda(\mathbf{y}_i-z),\mathbf{y}_{i+1:n}) \\
& \quad \left(\frac{1}{(s-t)^i} + \frac{v^{\frac{1}{2}}}{(s-t)^{2i-\frac{1}{2}}} \right) (s-t)^{(i-\frac{3}{2})(1+\eta)} \\
& \leq C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z-\mathbf{y}_i)}{v^{\frac{1}{2}}} \int_0^1 d\lambda \hat{q}_c(t,s,\mathbf{x},\mathbf{y}_{1:i-1},\mathbf{y}_i+\lambda(\mathbf{y}_i-z),\mathbf{y}_{i+1:n}) \\
& \quad \left(\frac{1}{(s-t)^{\frac{3}{2}-\eta(i-\frac{3}{2})}} + \frac{v^{\frac{1}{2}}}{(s-t)^{i-\frac{1}{2}+\frac{3}{2}-\eta(i-\frac{3}{2})}} \right).
\end{aligned}$$

Finally,

$$\begin{aligned}
& \int_0^{(s-t)^{\rho_{i,i}}} dv v^{\frac{\alpha_i^i-1}{2}} \int_{\mathbb{R}^d} dz |(\mathcal{F}_{2,i,(i,1),i}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}),(\mathbf{t},\mathbf{x})}| \\
& \leq C (s-t)^{-\frac{3}{2}+\eta(i-\frac{3}{2})} \hat{q}_{c \setminus i}(t,s,\mathbf{x},\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}) \int_0^{(s-t)^{\rho_{i,i}}} dv \left(v^{-1+\frac{\alpha_i^i}{2}} + \frac{v^{-\frac{1}{2}+\frac{\alpha_i^i}{2}}}{(s-t)^{i-\frac{1}{2}}} \right) \\
& \leq C (s-t)^{-1+\eta(i-\frac{3}{2})+\frac{\eta}{8}} \hat{q}_{c \setminus i}(t,s,\mathbf{x},\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}) \leq C (s-t)^{-\frac{3}{2}+\gamma_i^i} \hat{q}_{c \setminus i}(t,s,\mathbf{x},\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}),
\end{aligned}$$

recalling for the last inequality that, since from (3.13) and (3.14), $\alpha_i^i = (1+\eta/4)/(2i-1)$, $\rho_{i,i} = 2i-1$, $\alpha_i^i \rho_{i,i}/2 = 1/2 + \eta/8$, $\rho_{i,i}/2 = i-1/2$. We also refer to (3.15) for similar computations.

On the other hand, we readily get, similarly to the previous contributions from Lemmas 7 and 8:

$$\begin{aligned}
|\mathcal{F}_{3,i,(i,1)}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}),(\mathbf{t},\mathbf{x})}(v,z)| & \leq C \int_{\mathbb{R}^d} d\mathbf{y}_i \frac{h_{cv}(z-\mathbf{y}_i)}{v^{\frac{1}{2}}} \frac{|(\mathbf{y}-\boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{i-1}|^{1+\eta}}{(s-t)^i} \hat{q}_c(t,s,\mathbf{x},\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n}) \\
& \leq C (s-t)^{-\frac{3}{2}+\eta(i-\frac{3}{2})} v^{-\frac{1}{2}} \int_{\mathbb{R}^d} d\mathbf{y}_i h_{cv}(z-\mathbf{y}_i) \hat{q}_c(t,s,\mathbf{x},\mathbf{y}_{1:i-1},z,\mathbf{y}_{i+1:n}).
\end{aligned}$$

Similarly to the contributions appearing in (3.6) this bound needs to be integrated w.r.t. $v \in [0, (s-t)^{\rho_{i,i}}]$. One gets:

$$\begin{aligned}
& \int_0^{(s-t)^{\rho_{i,i}}} dv v^{\frac{\alpha_i^i-1}{2}} \int_{\mathbb{R}^d} dz |(\mathcal{F}_{3,i,(i,1)}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}),(\mathbf{t},\mathbf{x})}| \\
& \leq C (s-t)^{-\frac{3}{2}+\eta(i-\frac{3}{2})} \hat{q}_{c \setminus i}(t,s,\mathbf{x},\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}) \int_0^{(s-t)^{\rho_{i,i}}} dv v^{-1+\frac{\alpha_i^i}{2}} \\
& \leq C (s-t)^{-\frac{3}{2}+\eta(i-\frac{3}{2})+\frac{\alpha_i^i}{2}\rho_{i,i}} \hat{q}_{c \setminus i}(t,s,\mathbf{x},\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}) \\
& \leq C (s-t)^{-\frac{3}{2}+\gamma_i^i} \hat{q}_{c \setminus i}(t,s,\mathbf{x},\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}),
\end{aligned}$$

reproducing the previous computations done on $\mathcal{F}_{2,i,(i,1)}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}),(\mathbf{t},\mathbf{x})}(v,z)$ for the last inequality.

3.2 Proof of Lemma 10

We now tackle the Besov estimate of the cross derivative of the solution of (1.6). Fix $t \in [0, T]$ and $(\mathbf{x}_{1:i-1}, \mathbf{x}_{i+1:n}) \in \mathbb{R}^{(n-1)d}$. From the thermic characterization of Besov spaces recalled in equation (C.1), we actually have to control:

$$\begin{aligned} \|D_i D_k u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})\|_{B_{\infty, \infty}^{\alpha_i^k - 1}} &:= \|\varphi(D) D_i D_k u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})\|_{\infty} \\ &+ \sup_{v \in [0, 1]} v^{\frac{1-\alpha_i^k}{2}} \sup_{z \in \mathbb{R}^d} |h_v \star D_i D_k u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z)|. \end{aligned} \quad (3.19)$$

For the proof we focus on the second contribution in the above definition. The first one could be handled similarly and more directly⁹.

Using representation (2.20) we first write with the notations of (2.21):

$$\begin{aligned} &|h_v \star D_i D_k u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z)| \tag{3.20} \\ &= \left| \int_{\mathbb{R}^d} d\mathbf{x}_i h_v(z - \mathbf{x}_i) D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \mathbf{x}_i, \mathbf{x}_{i+1:n}) \right| \\ &= \left| \int_{\mathbb{R}^d} d\mathbf{x}_i D_{\mathbf{x}_i} h_v(z - \mathbf{x}_i) \otimes \left(D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \mathbf{x}_i, \mathbf{x}_{i+1:n}) - D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, z, \mathbf{x}_{i+1:n}) \right) \right|, \end{aligned}$$

integrating by parts and using a usual cancellation argument in the above equality.

To handle such a term, we are led to compare the difference of the gradients through their expansion (2.21). Two cases then arise: the system is globally in the *off-diagonal* regime, i.e. the spatial points \mathbf{x}_i and z are *far* w.r.t. their corresponding time scale (there exists c_0 such that $c_0 |\mathbf{x}_i - z| \geq (T - t)^{i-1/2}$ or equivalently $c_0 |\mathbf{x}_i - z|^{2/(2i-1)} \geq (T - t)$); the system is globally in the *diagonal* regime, i.e. the spatial points \mathbf{x}_i and z are *close* w.r.t. their corresponding time scale ($c_0 |\mathbf{x}_i - z|^{2/(2i-1)} < (T - t)$).

Since in the global *off-diagonal* regime the spatial points are *far*, it is not expectable to control suitably the expansion of the gradients around their difference. In this case, it is in fact more natural to expand each gradient term thanks to (2.21) taking as freezing point the associated spatial argument, i.e. $\boldsymbol{\xi} = \mathbf{x}$ for the first gradient and denoting by $\mathbf{x}' = (\mathbf{x}_{1:i-1}, z, \mathbf{x}_{i+1:n})$, $\boldsymbol{\xi}' = \mathbf{x}'$ for the second one. This allows to take advantage of the underlying smoothing properties in time of the gradient (cf. Section 3.2.1 below¹⁰).

On the other hand, in the *global diagonal* regime (when $c_0 |\mathbf{x}_i - z|^{2/(2i-1)} \leq (T - t)$), we are again faced with a regime dichotomy. Note indeed that, expanding the gradients in (3.20) from (2.21), we have to deal with a time integral associated with the source and perturbative terms. We are hence again led to separate, within this time integral, the *local off-diagonal* and *diagonal* regimes, i.e. w.r.t. the time integration variable s . We hence introduce the time set $\mathcal{S}_i = \{s \in [t, T] : (s - t) \leq c_0 |\mathbf{x}_i - z|^{2/(2i-1)}\}$ (for the same previous constant c_0) which corresponds to the *local off-diagonal* regime and the complementary set $\mathcal{S}_i^c = \{s \in [t, T] : (s - t) > c_0 |\mathbf{x}_i - z|^{2/(2i-1)}\}$ which corresponds to the *local diagonal* one.

As above, in the *local off-diagonal* regime, we will not expand the difference of the gradients and we will only use their underlying smoothing properties in time, working thus with their expansion around two different freezing points associated with the corresponding spatial arguments, as suggested by (2.21) (see again Section 3.2.1).

Concerning the *local diagonal* regime, the proximity of the spatial points suggests to expand the gradients through a Taylor expansion. Starting again from their corresponding representation of (2.21), it is natural to consider similar spatial freezing points. Such a strategy indeed yields to only consider spatial sensitivities of the underlying Gaussian proxy (see Section 3.2.2). Observe that keeping the two distinct freezing points would lead to investigate the full sensitivity between two different proxys, including the sensitivity of the corresponding covariance matrix and generator. Such an investigation appears to be quite involved. Furthermore, we did not succeed to make it work.

With our approach, we are led to expand one of the gradients in (3.20) around two different freezing points. Such a strategy was already used in the companion paper [CdRHM18] and leads to consider an additional boundary term arising precisely from the change of freezing point (see Section 3.2.3). Namely, we will expand the term $D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \mathbf{x}_i, \mathbf{x}_{i+1:n})$ with (2.21) taking $\boldsymbol{\xi} = \mathbf{x}$, whereas we will expand differently the contribution $D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, z, \mathbf{x}_{i+1:n})$, depending on the considered (local) regime (off-diagonal or diagonal)

⁹We could also observe from Proposition 14 that this term could also be directly bounded from the supremum norm of the gradient.

¹⁰note that in this case we have that for all s in $(t, T]$ the off-diagonal regime holds.

for the current running time. With the previous notations, this term will be expanded as in (2.21) around the freezing point $\xi' = \mathbf{x}'$ in the local off-diagonal regime and around $\tilde{\xi}' = \mathbf{x}$ in the local diagonal one. Denoting by $t_0 = t + c_0 |\mathbf{x}_i - z|^{2/(2i-1)}$ the transition time between the two regimes, we actually have:

$$\begin{aligned} u(t, \mathbf{x}') &= \int_t^T ds \left(\mathbb{I}_{\mathcal{S}_i} \left[\tilde{P}_{s,t}^{\xi'} f(s, \cdot) \right] (\mathbf{x}') + \mathbb{I}_{\mathcal{S}_i^c} \left[\tilde{P}_{s,t}^{\tilde{\xi}'} f(s, \cdot) \right] (\mathbf{x}') \right) \\ &\quad + \left[\tilde{P}_{t_0,t}^{\xi'} u(t_0, \cdot) \right] (\mathbf{x}') - \left[\tilde{P}_{t_0,t}^{\tilde{\xi}'} u(t_0, \cdot) \right] (\mathbf{x}') \\ &\quad + \int_t^T ds \left(\mathbb{I}_{\mathcal{S}_i} \left[\tilde{P}_{s,t}^{\xi'} \left((L_s - \tilde{L}_s^{\xi'}) u \right) (s, \cdot) \right] (\mathbf{x}') + \mathbb{I}_{\mathcal{S}_i^c} \left[\tilde{P}_{s,t}^{\tilde{\xi}'} \left((L_s - \tilde{L}_s^{\tilde{\xi}'}) u \right) (s, \cdot) \right] (\mathbf{x}') \right). \end{aligned} \quad (3.21)$$

We refer to Appendix B below for a proof of expansion (3.21) (see also Section 2.4 of the Detailed guide to the proof in [CDRHM18]). We again emphasize that in comparison with (2.21), the term in the second line in the r.h.s. of the above equation is the price to pay to consider different freezing points associated with the corresponding local off-diagonal and diagonal regimes.

From now on, we will assume w.l.o.g. that c_0 is a constant meant to be *small* (see Section 3.2.3, Lemma 12 and its proof). We also suppose that $|\mathbf{x}' - \mathbf{x}| \leq 1$ since otherwise the *global off-diagonal* regime holds, and the analysis of Section 3.2.1 applies.

Starting from (3.20), we expand $D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \mathbf{x}_i, \mathbf{x}_{i+1:n})$ using (2.21) with $\xi = \mathbf{x}$ and $D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, z, \mathbf{x}_{i+1:n})$ differentiating (3.21) w.r.t. \mathbf{x}_k and setting then $\xi' = \mathbf{x}'$, $\tilde{\xi}' = \mathbf{x}$. We rewrite:

$$\begin{aligned} h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z) &=: (h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z))|_{\mathcal{S}_i} \\ &\quad + (h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z))|_{\mathcal{S}_i^c} + (h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z))|_{\partial \mathcal{S}_i}, \end{aligned} \quad (3.22)$$

where,

$$\begin{aligned} &(h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z))|_{\mathcal{S}_i} \\ &:= \int_{\mathcal{S}_i} ds \int_{\mathbb{R}^d} d\mathbf{x}_i D_z h_v(z - \mathbf{x}_i) \left\{ \left[H_k^\xi(s, \mathbf{x}) - H_k^{\xi'}(s, \mathbf{x}') \right] + \left[I_k^\xi(s, \mathbf{x}) - I_k^{\xi'}(s, \mathbf{x}') \right] \right\}, \end{aligned} \quad (3.23)$$

corresponds to the difference of the previous expansions on the off-diagonal regime,

$$\begin{aligned} &(h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z))|_{\mathcal{S}_i^c} \\ &:= \int_{\mathcal{S}_i^c} ds \int_{\mathbb{R}^d} d\mathbf{x}_i D_z h_v(z - \mathbf{x}_i) \left\{ \left[H_k^\xi(s, \mathbf{x}) - H_k^{\tilde{\xi}'}(s, \mathbf{x}') \right] + \left[I_k^\xi(s, \mathbf{x}) - I_k^{\tilde{\xi}'}(s, \mathbf{x}') \right] \right\}, \end{aligned} \quad (3.24)$$

is the contribution of the diagonal regime and

$$\begin{aligned} &(h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z))|_{\partial \mathcal{S}_i} \\ &:= \int_{\mathbb{R}^d} d\mathbf{x}_i D_z h_v(z - \mathbf{x}_i) \left\{ D_{\mathbf{x}_k} \tilde{P}_{t_0,t}^{\xi'} u(t_0, \mathbf{x}') - D_{\mathbf{x}_k} \tilde{P}_{t_0,t}^{\tilde{\xi}'} u(t_0, \mathbf{x}') \right\}, \end{aligned} \quad (3.25)$$

is the resulting boundary term. This last term, arising from the change of freezing point, is particularly delicate to analyze.

3.2.1 Off-diagonal estimates: control of (3.23).

On the time set \mathcal{S}_i , we cannot expect some regularization from the difference of the transition densities so that we bluntly estimate the terms appearing in (3.23), writing:

$$\begin{aligned}
& \left| h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z) \right|_{\mathcal{S}_i} \\
& := \left| \int_{\mathbb{R}^d} d\mathbf{x}_i D_z h_v(z - \mathbf{x}_i) \left(\int_{\mathcal{S}_i} ds \left[H_k^\xi(s, \mathbf{x}) - H_k^{\xi'}(s, \mathbf{x}') \right] \right. \right. \\
& \quad \left. \left. + \int_{\mathcal{S}_i} ds \left[I_k^\xi(s, \mathbf{x}) - I_k^{\xi'}(s, \mathbf{x}') \right] \right) \right|_{(\xi, \xi')=(\mathbf{x}, \mathbf{x}')} \\
& \leq \int_{\mathbb{R}^d} d\mathbf{x}_i \frac{h_{cv}(z - \mathbf{x}_i)}{v^{\frac{1}{2}}} \int_{\mathcal{S}_i} ds \left(\left| H_k^\xi(s, \mathbf{x}) \right|_{\xi=\mathbf{x}} + \left| H_k^{\xi'}(s, \mathbf{x}') \right|_{\xi'=\mathbf{x}'} \right. \\
& \quad \left. + \left| I_k^\xi(s, \mathbf{x}) \right|_{\xi=\mathbf{x}} + \left| I_k^{\xi'}(s, \mathbf{x}') \right|_{\xi'=\mathbf{x}'} \right).
\end{aligned} \tag{3.26}$$

Those terms can then be handled following the previous analysis performed in Theorem 3 and Lemma 9, observing here that, w.r.t. the previous proofs, the above terms are not differentiated w.r.t. \mathbf{x}_1 . This improves the exponents of the time singularities of $1/2$. Similarly to (2.23), (2.25), (2.27) and (2.28) this therefore yields for the terms $H_k^\xi(s, \mathbf{x})$:

$$\left| H_k^\xi(s, \mathbf{x}) \right|_{\xi=\mathbf{x}} \leq C((s-t)^{-1+\delta_i^k} + \|\mathbf{D}D_1 u\|_\infty), \tag{3.27}$$

with $1 > \delta_i^k > 1/2$.

Reproducing the arguments that led to equations (2.31), (2.32), (2.39) and the statement of Lemma 9, exploiting again that there is now no differentiation w.r.t. \mathbf{x}_1 we get with the notations of (2.35)

$$\left| I_k^\xi(s, \mathbf{x}) \right|_{\xi=\mathbf{x}} \leq C(s-t)^{-1+\gamma_i^k} (\|\mathbf{D}u\|_\infty + \sup_{\bar{s} \in [0, T]} \|D_i D_k u^i(s, \cdot)\|_{B_{\infty, \infty}^{\alpha_i^k - 1}}).$$

Similar bounds hold for $\left| H_k^{\xi'}(s, \mathbf{x}') \right|_{\xi'=\mathbf{x}'}$ and $\left| I_k^{\xi'}(s, \mathbf{x}') \right|_{\xi'=\mathbf{x}'}$. Hence, from (3.27) and (3.28),

$$\begin{aligned}
& \int_{\mathcal{S}_i} ds \left(\left| H_k^\xi(s, \mathbf{x}) \right|_{\xi=\mathbf{x}} + \left| H_k^{\xi'}(s, \mathbf{x}') \right|_{\xi'=\mathbf{x}'} + \left| I_k^\xi(s, \mathbf{x}) \right|_{\xi=\mathbf{x}} + \left| I_k^{\xi'}(s, \mathbf{x}') \right|_{\xi'=\mathbf{x}'} \right) \\
& \leq C \left(\int_t^{(t+c_0|\mathbf{x}_i-z|^{\frac{2}{2i-1}}) \wedge T} ds \left((s-t)^{-1+\delta_i^k} \|\mathbf{D}u\|_\infty + \|\mathbf{D}D_1 u\|_\infty \right) \right. \\
& \quad \left. + (s-t)^{-1+\gamma_i^k} (\|\mathbf{D}u\|_\infty + \sup_{\bar{s} \in [0, T]} \|(D_{\mathbf{x}_i} D_{\mathbf{x}_k} u)_i(\bar{s}, \cdot)\|_{B_{\infty, \infty}^{\alpha_i^k - 1}}) \right) \\
& \leq C \left(|\mathbf{x}_i - z|^{\frac{2\delta_i^k}{2i-1}} \wedge (T-t)^{\delta_i^k} (\|\mathbf{D}u\|_\infty + \|\mathbf{D}D_1 u\|_\infty) \right. \\
& \quad \left. + |\mathbf{x}_i - z|^{\frac{2\gamma_i^k}{2i-1}} \wedge (T-t)^{\gamma_i^k} (\|\mathbf{D}u\|_\infty + \sup_{\bar{s} \in [0, T]} \|(D_{\mathbf{x}_i} D_{\mathbf{x}_k} u)_i(\bar{s}, \cdot)\|_{B_{\infty, \infty}^{\alpha_i^k - 1}}) \right) \\
& \leq C |\mathbf{x}_i - z|^{\alpha_i^k} T^{\delta'} \left(\|\mathbf{D}u\|_\infty + \|\mathbf{D}D_1 u\|_\infty + \sup_{\bar{s} \in [0, T]} \|(D_{\mathbf{x}_i} D_{\mathbf{x}_k} u)_i(\bar{s}, \cdot)\|_{B_{\infty, \infty}^{\alpha_i^k - 1}} \right),
\end{aligned} \tag{3.28}$$

for some $\delta' := \delta'(\mathbf{A}) > 0$, denoting by

$$\sup_{\bar{s} \in [0, T]} \|(D_{\mathbf{x}_i} D_{\mathbf{x}_k} u)_i(\bar{s}, \cdot)\|_{B_{\infty, \infty}^{\alpha_i^k - 1}} := \sup_{\bar{s} \in [0, T], \mathbf{z}_j, j \in \llbracket 1, n \rrbracket, j \neq i} \|D_{\mathbf{x}_i} D_{\mathbf{x}_k} u^i(\bar{s}, \mathbf{z}_{1:i-1}, \cdot, \mathbf{z}_{i+1:n})\|_{B_{\infty, \infty}^{\alpha_i^k - 1}}, \tag{3.29}$$

and recalling for the last inequality that $\gamma_i^k = 1/2 + \eta(i-3/2)$ so that $2\gamma_i^k/(2i-1) = [1+2\eta(i-3/2)]/(2i-1) > (1+\eta/4)/(2i-1) = \alpha_i^k$ (see also the statements of Lemmas 9 and 10) and similarly for the contributions involving

$\delta_i^k > 1/2$. We eventually get from (3.28) and (3.26):

$$\begin{aligned}
& |h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z)| \Big|_{\mathcal{S}_i} \\
& \leq CT^{\delta'} \left(\|\mathbf{D}u\|_\infty + \|\mathbf{D}D_1u\|_\infty + \sup_{\bar{s} \in [0, T]} \|(D_{\mathbf{x}_i} D_{\mathbf{x}_k} u)_i(\bar{s}, \cdot)\|_{B_{\infty, \infty}^{\alpha_i^k - 1}} \right) \int_{\mathbb{R}^d} d\mathbf{x}_i \frac{h_{cv}(z - \mathbf{x}_i)}{v^{\frac{1}{2}}} |z - \mathbf{x}_i|^{\alpha_i^k} \\
& \leq CT^{\delta'} \left(\|\mathbf{D}u\|_\infty + \|\mathbf{D}D_1u\|_\infty + \sup_{\bar{s} \in [0, T]} \|(D_{\mathbf{x}_i} D_{\mathbf{x}_k} u)_i(\bar{s}, \cdot)\|_{B_{\infty, \infty}^{\alpha_i^k - 1}} \right) v^{\frac{\alpha_i^k - 1}{2}}. \tag{3.30}
\end{aligned}$$

3.2.2 Diagonal estimates: control of the term (3.24).

We are now going to handle the term (3.24) which correspond to the r.h.s. of (3.20) on \mathcal{S}_i^c . In that case the points \mathbf{x}_i and \mathbf{x}'_i are close w.r.t. the characteristic time scale of the i^{th} variable and the main idea consists in controlling the difference between the frozen densities at $\boldsymbol{\xi} = \tilde{\boldsymbol{\xi}}' = \mathbf{x}$ with starting points \mathbf{x} and \mathbf{x}' respectively. Precisely, recalling that \mathbf{x} and \mathbf{x}' only differ in the i^{th} component, we can write:

$$D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}', \mathbf{y}) = - \int_0^1 d\lambda D_{\mathbf{x}_i} \left(D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x} + \lambda(\mathbf{x}' - \mathbf{x}), \mathbf{y}) \right) \cdot (\mathbf{x}' - \mathbf{x})_i. \tag{3.31}$$

From Lemma 7 we thus derive:

$$|D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}', \mathbf{y})| \leq \frac{C|(\mathbf{x}' - \mathbf{x})_i|}{(s-t)^{(i-\frac{1}{2})+(k-\frac{1}{2})}} \int_0^1 d\lambda \hat{p}_{C-1}^\xi(t, s, \mathbf{x} + \lambda(\mathbf{x}' - \mathbf{x}), \mathbf{y}). \tag{3.32}$$

Now, from the definition of \hat{p}_{C-1} in Proposition 5, recalling as well from (2.7) that $\mathbf{x} \mapsto \mathbf{m}_{s,t}^\xi(\mathbf{x})$ is affine, we get:

$$\begin{aligned}
& |\hat{p}_{C-1}^\xi(t, s, \mathbf{x} + \lambda(\mathbf{x}' - \mathbf{x}), \mathbf{y})| \\
& \leq \frac{C}{(s-t)^{\frac{n^2 d}{2}}} \exp(-c(s-t)|\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^\xi(\mathbf{x} + \lambda(\mathbf{x}' - \mathbf{x})) - \mathbf{y})|^2) \\
& \leq \frac{C}{(s-t)^{\frac{n^2 d}{2}}} \exp(c(s-t)|\mathbb{T}_{s-t}^{-1}\mathbf{m}_{s,t}^\xi(\mathbf{x} - \mathbf{x}')|^2) \exp(-\frac{c}{2}(s-t)|\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^\xi(\mathbf{x}) - \mathbf{y})|^2).
\end{aligned}$$

Using the rescaling arguments of the proof of Proposition 5 on the resolvent (see equation (2.13)), we then get $(s-t)^{1/2}|\mathbb{T}_{s-t}^{-1}\mathbf{m}_{s,t}^\xi(\mathbf{x} - \mathbf{x}')| \leq C(s-t)^{1/2}|\mathbb{T}_{s-t}^{-1}(\mathbf{x} - \mathbf{x}')| = C(s-t)^{-i+1/2}|(\mathbf{x}' - \mathbf{x})_i| \leq C$, from the very definition of \mathcal{S}_i^c . Hence,

$$|\hat{p}_{C-1}^\xi(t, s, \mathbf{x} + \lambda(\mathbf{x}' - \mathbf{x}), \mathbf{y})| \leq \frac{C}{(s-t)^{\frac{n^2 d}{2}}} \exp(-\frac{c}{2}(s-t)|\mathbb{T}_{s-t}^{-1}(\mathbf{m}_{s,t}^\xi(\mathbf{x}) - \mathbf{y})|^2),$$

so that, from (3.32) and recalling that on \mathcal{S}_i^c , $|(\mathbf{x}' - \mathbf{x})_i|/(s-t) \leq C(|(\mathbf{x}' - \mathbf{x})_i|/(s-t))^{\alpha_i^k}$, the following important control holds:

$$|D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}', \mathbf{y})| \leq \frac{C|(\mathbf{x}' - \mathbf{x})_i|^{\alpha_i^k}}{(s-t)^{\alpha_i^k(i-\frac{1}{2})+(k-\frac{1}{2})}} \hat{p}_{C-1}^\xi(t, s, \mathbf{x}, \mathbf{y}). \tag{3.33}$$

Write now from (3.24), recalling that $\tilde{\boldsymbol{\xi}}' = \boldsymbol{\xi}$,

$$\begin{aligned}
& |h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z)| \Big|_{\mathcal{S}_i^c} \\
& := \left| \int_{\mathbb{R}^d} d\mathbf{x}_i D_z h_v(z - \mathbf{x}_i) \left(\int_{\mathcal{S}_i^c} ds [H_k^\xi(s, \mathbf{x}) - H_k^\xi(s, \mathbf{x}')] + \int_{\mathcal{S}_i^c} ds [I_k^\xi(s, \mathbf{x}) - I_k^\xi(s, \mathbf{x}')] \right) \right|, \tag{3.34}
\end{aligned}$$

and let us discuss how the terms $H_k^\xi(s, \mathbf{x}) - H_k^\xi(s, \mathbf{x}')$, $I_k^\xi(s, \mathbf{x}) - I_k^\xi(s, \mathbf{x}')$ in the above equation can be handled.

We first focus on the term $I_k^\xi(s, \mathbf{x}) - I_k^\xi(s, \mathbf{x}')$ in (3.34). This contribution, associated with the degenerate components of perturbed operator, is again the most delicate to handle. From the definitions in (2.30) we are

led to control the sum $\sum_{\ell=1}^3 [I_{\ell,k}^{\xi}(s, \mathbf{x}) - I_{\ell,k}^{\xi}(s, \mathbf{x}')]$. For the terms $I_{1,k}^{\xi}(s, \mathbf{x}) - I_{1,k}^{\xi}(s, \mathbf{x}')$, $I_{2,k}^{\xi}(s, \mathbf{x}) - I_{2,k}^{\xi}(s, \mathbf{x}')$ we are going to reproduce the analysis leading to (2.31), (2.32). Observe first that the above terms do not involve $D_{\mathbf{x}_1}$, therefore we gain a singularity of order $1/2$ w.r.t. to the indicated equations (2.31), (2.32). On the other hand, the difference of the derivatives of the frozen densities w.r.t. \mathbf{x}_k can be handled with (3.33). This leads to:

$$\begin{aligned} & \left| \sum_{\ell=1}^2 \int_{t+c_0|\mathbf{x}'-\mathbf{x}|_i}^{T} \frac{2}{2^i-1} ds \left(I_{\ell,k}^{\xi}(s, \mathbf{x}) - I_{\ell,k}^{\xi}(s, \mathbf{x}') \right) \right| \\ & \leq C |\mathbf{x}' - \mathbf{x}|_i^{\alpha_i^k} \|\mathbf{D}u\|_{\infty} \sum_{j=k}^n \int_t^T ds (s-t)^{-(k-\frac{1}{2})-\alpha_i^k(i-\frac{1}{2})} \left((s-t)^{\beta_j(j-\frac{1}{2})} + (s-t)^{(1+\eta)(j-\frac{1}{2})} \right), \end{aligned}$$

changing the summation variables from (2.31) for notational simplicity.

From the very definition of $\alpha_i^k = (1 + \eta/4)/(2i - 1)$ in Lemma 9 and the specific choice of η in assumption **(A)** (see **(H $_{\eta}$)**) we derive

$$\left| \sum_{\ell=1}^2 \int_{t+c_0|\mathbf{x}'-\mathbf{x}|_i}^{T} \frac{2}{2^i-1} ds \left(I_{\ell,k}^{\xi}(s, \mathbf{x}) - I_{\ell,k}^{\xi}(s, \mathbf{x}') \right) \right| \leq C |\mathbf{x}' - \mathbf{x}|_i^{\alpha_i^k} \|\mathbf{D}u\|_{\infty} T^{\delta}, \quad (3.35)$$

for some $\delta > 0$.

From the previous analysis it is therefore sufficient to focus on the tricky term, namely $I_{3,k}(s, \mathbf{x})$ introduced in (2.30). We begin the proof considering first $I_{3,k}(s, \mathbf{x})$. Exploiting as well Lemma 7 for a centering argument w.r.t. the k^{th} variable, we write:

$$\begin{aligned} & I_{3,k}^{\xi}(s, \mathbf{x}) - I_{3,k}^{\xi}(s, \mathbf{x}') \\ & = \sum_{\ell=2}^k \int_{\mathbb{R}^{nd}} d\mathbf{y} \left(\mathbf{F}_{\ell}(s, \mathbf{y}_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})) - \mathbf{F}_{\ell}(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{\ell-1} \mathbf{F}_{\ell}(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{\ell-1} \right) \\ & \quad \times \left(D_{\mathbf{y}_{\ell}} u(s, \mathbf{y}) - D_{\mathbf{y}_{\ell}} u(s, \mathbf{y}_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})) \right) \left(D_{\mathbf{x}_k} \tilde{p}^{\xi}(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_k} \tilde{p}^{\xi}(t, s, \mathbf{x}', \mathbf{y}) \right). \end{aligned}$$

Let us reproduce now the arguments used in Section 2.2 to handle $I_{3,k}^{\xi}$ (see e.g. the computations from equation (2.33) to (2.34)). Expanding with the Taylor formula the difference $\left(D_{\mathbf{y}_{\ell}} u(s, \mathbf{y}) - D_{\mathbf{y}_{\ell}} u(s, \mathbf{y}_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})) \right)$, using the Schwarz theorem to exchange the order of differentiations¹¹, we obtain with the notations of (2.35) (see also the comments following this equation):

$$\begin{aligned} & I_{3,k}^{\xi}(s, \mathbf{x}) - I_{3,k}^{\xi}(s, \mathbf{x}') \\ & = \sum_{m=k}^n \sum_{\ell=2}^k \int_0^1 d\lambda \int_{\mathbb{R}^{nd}} d\mathbf{y} \\ & \quad \left[\left(\mathbf{F}_{\ell}(s, \mathbf{y}_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})) - \mathbf{F}_{\ell}(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{\ell-1} \mathbf{F}_{\ell}(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{\ell-1} \right) \right. \\ & \quad \left. \times \left(D_{\mathbf{x}_k} \tilde{p}^{\xi}(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_k} \tilde{p}^{\xi}(t, s, \mathbf{x}', \mathbf{y}) \right) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_m \right] \\ & \quad D_{\mathbf{y}_{\ell}} D_{\mathbf{y}_m} u(s, \mathbf{y}_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})) + \lambda (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{k:n} \\ & =: \sum_{m=k}^n \sum_{\ell=2}^k \int_0^1 d\lambda \int_{\mathbb{R}^{nd}} d\mathbf{y} \left[\Psi_{\ell,k,m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x})}(\mathbf{y}_{\ell}) - \Psi_{\ell,k,m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x}')}(\mathbf{y}_{\ell}) \right] \\ & \quad D_{\mathbf{y}_{\ell}} D_{\mathbf{y}_m} u(s, \mathbf{y}_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})) + \lambda (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{k:n}. \end{aligned}$$

¹¹Recall indeed that what we are able to control is precisely the Hölder moduli of the derivatives $D_{\mathbf{y}_m} u(s, \cdot)$ w.r.t. variables $\ell \leq m$.

Thus, we derive similarly to (2.39):

$$\begin{aligned}
|I_{3,k}^\xi(s, \mathbf{x}) - I_{3,k}^\xi(s, \mathbf{x}')| &\leq \sum_{m=k}^n \sum_{\ell=2}^k \int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}) \\
&\quad \left\{ \left\| \Psi_{\ell,k,m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x})}(\cdot) - \Psi_{\ell,k,m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x}')}(\cdot) \right\|_{B_{1,1}^{1-\alpha_\ell^m}} \right. \\
&\quad \left. \times \sup_{\mathbf{z}_j, j \in [1, n], j \neq \ell} \left\| D_\ell D_m u(s, \mathbf{z}_{1:\ell-1}, \cdot, \mathbf{z}_{\ell+1:n}) \right\|_{B_{\infty, \infty}^{\alpha_\ell^m - 1}} \right\},
\end{aligned} \tag{3.36}$$

where $\int_{\mathbb{R}^{(n-1)d}} d(\mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n})$ means that we integrate over $\mathbf{y}_{1:\ell-1}$ and $\mathbf{y}_{\ell+1:n}$. To conclude, we need the following appropriate version of Lemma 9 to handle the Besov norm with negative exponent in the above r.h.s. Its proof is postponed to the next section.

Lemma 11. *Let $k \in \llbracket 2, n \rrbracket$, $\ell \in \llbracket 2, k \rrbracket$ and $m \in \llbracket k, n \rrbracket$ and let $\Psi_{\ell, (k, i), m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x})} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ be the function defined by (2.35). There exist $C := C(\mathbf{A}) > 0$ and $\gamma_\ell^m := \gamma_\ell^m(\mathbf{A}) := 1/2 + \eta(\ell - 3/2) > 1/2$ such that*

$$\begin{aligned}
&\left\| \Psi_{\ell, k, m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x})}(\cdot) - \Psi_{\ell, k, m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x}')}(\cdot) \right\|_{B_{1,1}^{1-\alpha_\ell^m}} \\
&\leq C \hat{q}_{c \setminus \ell}(t, s, \mathbf{x}, (\mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n})) (s-t)^{-1-(i-\frac{1}{2})\alpha_i^k + \gamma_\ell^m} |(\mathbf{x} - \mathbf{x}')_i|^{\alpha_i^k},
\end{aligned}$$

with $\hat{q}_{c \setminus \ell}(t, s, \mathbf{x}, (\mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}))$ as in (2.40).

Again, for the specific choice of $\alpha_i^k = (1 + \eta/4)/(2i - 1)$ performed in the proof of Lemma 9, we eventually derive from Lemma 11 and (3.36), with the notation of (3.29), that:

$$\begin{aligned}
&\int_{\mathcal{S}_i^c} ds |I_{3,k}^\xi(s, \mathbf{x}) - I_{3,k}^\xi(s, \mathbf{x}')| \\
&\leq C \left(\sum_{m=k}^n \sum_{\ell=2}^k \sup_{s \in [0, T]} \|(D_\ell D_m u)_\ell(s, \cdot)\|_{B_{\infty, \infty}^{-1+\alpha_\ell^m}} \int_t^T ds (s-t)^{-1-(i-\frac{1}{2})\alpha_i^k + \gamma_\ell^m} \right) |(\mathbf{x} - \mathbf{x}')_i|^{\alpha_i^k} \\
&\leq CT^\delta \sup_{m \in \llbracket k, n \rrbracket, \ell \in \llbracket 1, k \rrbracket, s \in [0, T]} \|(D_\ell D_m u)_\ell(s, \cdot)\|_{B_{\infty, \infty}^{\alpha_\ell^m - 1}} |(\mathbf{x} - \mathbf{x}')_i|^{\alpha_i^k},
\end{aligned} \tag{3.37}$$

for some $\delta := 3\eta/8 > 0$. Combining this estimate together with (3.35) we eventually derive

$$\begin{aligned}
&\left| \int_{\mathcal{S}_i^c} ds \left[I_k^\xi(s, \mathbf{x}) - I_k^\xi(s, \mathbf{x}') \right] \right| \\
&\leq T^\delta C |(\mathbf{x}' - \mathbf{x})_i|^{\alpha_i^k} \left(\|\mathbf{D}u\|_\infty + \sup_{m \in \llbracket k, n \rrbracket, \ell \in \llbracket 1, k \rrbracket, s \in [0, T]} \|(D_\ell D_m u)_\ell(s, \cdot)\|_{B_{\infty, \infty}^{\alpha_\ell^m - 1}} \right).
\end{aligned} \tag{3.38}$$

Now, the term $H_k^\xi(s, \mathbf{x}) - H_k^\xi(s, \mathbf{x}')$ in (3.34) (non-degenerate variables) can be handled reproducing the same previous arguments for $I_k^\xi(s, \mathbf{x}) - I_k^\xi(s, \mathbf{x}')$, exploiting (3.33) and following the computations performed for H_k in the proof of Theorem 3 (see e.g. (2.29)). From the definition of \mathcal{S}_i^c we obtain:

$$\begin{aligned}
&\left| \int_{t+c_0|(\mathbf{x}' - \mathbf{x})_i|^{\frac{2}{2i-1}}}^T ds \left(H_k^\xi(s, \mathbf{x}) - H_k^\xi(s, \mathbf{x}') \right) \right| \\
&\leq C |(\mathbf{x}' - \mathbf{x})_i|^{\alpha_i^k} (\|\mathbf{D}D_{\mathbf{x}_1} u\|_\infty + \|\mathbf{D}u\|_\infty) \int_t^T ds (s-t)^{-1-\alpha_i^k(i-\frac{1}{2})+\delta_i^k} \\
&\leq C |(\mathbf{x}' - \mathbf{x})_i|^{\alpha_i^k} (\|\mathbf{D}D_{\mathbf{x}_1} u\|_\infty + \|\mathbf{D}u\|_\infty) T^\delta,
\end{aligned} \tag{3.39}$$

for some $\delta > 0$ recalling that $\delta_i^k > 1/2$ for the last inequality. The arguments needed to control this term are actually those already exploited in [CdR17] when $n = 2$.

Gathering equations (3.38) and (3.39), we finally derive with the notations of (3.34):

$$\begin{aligned}
& |h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z)| \Big|_{S_i^c} \\
& \leq CT^\delta \int_{\mathbb{R}^d} d\mathbf{x}_i \frac{h_{cv}(z - \mathbf{x}_i)}{v^{\frac{1}{2}}} |\mathbf{x}_i - z|^{\alpha_i^k} \left(\| \mathbf{D} D_{\mathbf{x}_1} u \|_\infty + \| \mathbf{D} u \|_\infty + \sup_{2 \leq \ell \leq m \leq n, s \in [0, T]} \| (D_\ell D_m u)_\ell(s, \cdot) \|_{B_{\infty, \infty}^{\alpha_\ell^{m-1}}} \right) \\
& \leq CT^\delta v^{\frac{\alpha_i^k - 1}{2}} \left(\| \mathbf{D} D_{\mathbf{x}_1} u \|_\infty + \| \mathbf{D} u \|_\infty + \sup_{2 \leq \ell \leq m \leq n, s \in [0, T]} \| (D_\ell D_m u)_\ell(s, \cdot) \|_{B_{\infty, \infty}^{\alpha_\ell^{m-1}}} \right). \tag{3.40}
\end{aligned}$$

3.2.3 Discontinuity term associated with the regime time change: control of the term (3.25).

We here aim at handling

$$\begin{aligned}
& (h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z)) \Big|_{\partial S_i} \\
& = \int_{\mathbb{R}^d} d\mathbf{x}_i D_z h_v(z - \mathbf{x}_i) \left\{ D_{\mathbf{x}_k} \tilde{P}_{t_0, t}^{\xi'} u(t_0, \mathbf{x}') - D_{\mathbf{x}_k} \tilde{P}_{t_0, t}^{\tilde{\xi}'} u(t_0, \mathbf{x}') \right\},
\end{aligned}$$

which we will actually handle like the off-diagonal components. Recall here that the transition time $t_0 = t + c_0 |\mathbf{x} - \mathbf{x}'|^{2/(2i-1)}$ and that $(\xi', \tilde{\xi}') = (\mathbf{x}', \mathbf{x})$, where these values are again plugged in the expansions after differentiating. From Lemma 7 (cancellation argument), we write:

$$\begin{aligned}
& D_{\mathbf{x}_k} \tilde{P}_{t_0, t}^{\xi'} u(t_0, \mathbf{x}') - D_{\mathbf{x}_k} \tilde{P}_{t_0, t}^{\tilde{\xi}'} u(t_0, \mathbf{x}') \\
& = \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) [u(t_0, \mathbf{y}) - u(t_0, \mathbf{y}_{1:k-1}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k:n})] d\mathbf{y} \\
& \quad - \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{\tilde{\xi}'}(t, t_0, \mathbf{x}', \mathbf{y}) [u(t_0, \mathbf{y}) - u(t_0, \mathbf{y}_{1:k-1}, (\mathbf{m}_{t_0, t}^{\xi'}(\mathbf{x}'))_k, \boldsymbol{\theta}_{t_0, t}(\mathbf{x}')_{k+1:n})] d\mathbf{y}.
\end{aligned}$$

We now split the above contribution into three terms:

$$D_{\mathbf{x}_k} \tilde{P}_{t_0, t}^{\xi'} u(t_0, \mathbf{x}') - D_{\mathbf{x}_k} \tilde{P}_{t_0, t}^{\tilde{\xi}'} u(t_0, \mathbf{x}') =: (\mathcal{B}_1^{\xi'; \tilde{\xi}'} + \mathcal{B}_2^{\xi'; \tilde{\xi}'} + \mathcal{B}_3^{\xi'; \tilde{\xi}'})(t_0, \mathbf{x}'), \tag{3.41}$$

where

$$\begin{aligned}
\mathcal{B}_1^{\xi'; \tilde{\xi}'}(t_0, \mathbf{x}') & = \left\{ \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) [u(t_0, \mathbf{y}) - u(t_0, \mathbf{y}_{1:k}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k+1:n})] d\mathbf{y} \right. \\
& \quad \left. - \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{\tilde{\xi}'}(t, t_0, \mathbf{x}', \mathbf{y}) [u(t_0, \mathbf{y}) - u(t_0, \mathbf{y}_{1:k}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k+1:n})] d\mathbf{y} \right\}, \tag{3.42}
\end{aligned}$$

$$\begin{aligned}
\mathcal{B}_2^{\xi'; \tilde{\xi}'}(t_0, \mathbf{x}') & := \left\{ \left[\int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \left[u(t_0, \mathbf{y}_{1:k}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k+1:n}) - u(t_0, \mathbf{y}_{1:k-1}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k:n}) \right. \right. \right. \\
& \quad \left. \left. - \langle D_k u(t_0, \mathbf{y}_{1:k-1}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k:n}), (\mathbf{y} - \boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_k \rangle \right] d\mathbf{y} \right] \\
& \quad - \left[\int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{\tilde{\xi}'}(t, t_0, \mathbf{x}', \mathbf{y}) \left[u(t_0, \mathbf{y}_{1:k}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k+1:n}) - u(t_0, \mathbf{y}_{1:k-1}, (\mathbf{m}_{t_0, t}^{\xi'}(\mathbf{x}'))_k, \boldsymbol{\theta}_{t_0, t}(\mathbf{x}')_{k+1:n}) \right. \right. \\
& \quad \left. \left. - \langle D_k u(t_0, \mathbf{y}_{1:k-1}, (\mathbf{m}_{t_0, t}^{\xi'}(\mathbf{x}'))_k, \boldsymbol{\theta}_{t_0, t}(\mathbf{x}')_{k+1:n}), (\mathbf{y} - \mathbf{m}_{t_0, t}^{\xi'}(\mathbf{x}'))_k \rangle \right] d\mathbf{y} \right] \right\}, \tag{3.43}
\end{aligned}$$

$$\begin{aligned}
\mathcal{B}_3^{\xi'; \tilde{\xi}'}(t_0, \mathbf{x}') & := \left\{ \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \langle D_k u(t_0, \mathbf{y}_{1:k-1}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k:n}), (\mathbf{y} - \boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_k \rangle d\mathbf{y} \right. \\
& \quad \left. - \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{\tilde{\xi}'}(t, t_0, \mathbf{x}', \mathbf{y}) \langle D_k u(t_0, \mathbf{y}_{1:k-1}, (\mathbf{m}_{t_0, t}^{\xi'}(\mathbf{x}'))_k, \boldsymbol{\theta}_{t_0, t}(\mathbf{x}')_{k+1:n}), (\mathbf{y} - \mathbf{m}_{t_0, t}^{\xi'}(\mathbf{x}'))_k \rangle d\mathbf{y} \right\} \tag{3.44}
\end{aligned}$$

We now exploit the Hölder regularity of $D_k u$ w.r.t. the k^{th} variable to control the terms in $\mathcal{B}_2^{\xi', \tilde{\xi}'}(t_0, \mathbf{x}')$ defined in (3.43). Let us first write from the previous decomposition:

$$\begin{aligned} |\mathcal{B}_{2,1}^{\xi', \tilde{\xi}'}(t_0, \mathbf{x}')| &:= \left| \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \left[u(t_0, \mathbf{y}_{1:k}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k+1:n}) - u(t_0, \mathbf{y}_{1:k-1}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k:n}) \right. \right. \\ &\quad \left. \left. - \langle D_k u(t_0, \mathbf{y}_{1:k-1}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k:n}), (\mathbf{y} - \boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_k \rangle \right] d\mathbf{y} \right| \\ &= \int_{\mathbb{R}^{nd}} \int_0^1 d\lambda \left| D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \left[D_k u(t_0, \mathbf{y}_{1:k-1}, \cdot, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k+1:n}) \right]_{\alpha_k^k} |(\mathbf{y} - \boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_k|^{1+\alpha_k^k} d\mathbf{y}, \right. \end{aligned}$$

where $[\cdot]_{\alpha_k^k}$ denotes the Hölder modulus of order α_k^k . From Proposition 5, we thus derive:

$$\begin{aligned} |\mathcal{B}_{2,1}^{\xi', \tilde{\xi}'}(t_0, \mathbf{x}')| &\leq C(t_0 - t)^{\alpha_k^k(k - \frac{1}{2})} \sup_{\mathbf{z}_j, j \in \llbracket 1, n \rrbracket, j \neq k} [D_k u(t_0, \mathbf{z}_{1:k}, \cdot, \mathbf{z}_{k+1:n})]_{\alpha_k^k} \\ &\quad \times \int_{\mathbb{R}^{(n-1)d}} d\mathbf{y}_{1:k-1} d\mathbf{y}_{k+1:n} \hat{q}_{C^{-1} \setminus k}(t, s, \mathbf{x}, (\mathbf{y}_{1:k-1}, \mathbf{y}_{k+1:n})) \\ &= C(t_0 - t)^{\frac{1}{2} + \frac{\eta}{8}} [(D_k u)_k(t_0, \cdot)]_{\alpha_k^k}, \end{aligned} \quad (3.45)$$

recalling from (3.13) that $\alpha_k^k = (1 + \eta/4)/(2k - 1)$ and using the notation of (1.3) for the last inequality. The same arguments readily give:

$$\begin{aligned} |\mathcal{B}_{2,2}^{\xi', \tilde{\xi}'}(t_0, \mathbf{x}')| &:= \left| \int_{\mathbb{R}^{nd}} D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \left[u(t_0, \mathbf{y}_{1:k}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k+1:n}) - u(t_0, \mathbf{y}_{1:k-1}, (\mathbf{m}_{t_0, t}^{\xi'}(\mathbf{x}'))_k, \boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k+1:n}) \right. \right. \\ &\quad \left. \left. - \langle D_k u(t_0, \mathbf{y}_{1:k-1}, (\mathbf{m}_{t_0, t}^{\xi'}(\mathbf{x}'))_k, \boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k+1:n}), (\mathbf{y} - \mathbf{m}_{t_0, t}^{\xi'}(\mathbf{x}'))_k \rangle \right] d\mathbf{y} \right| \\ &\leq C(t_0 - t)^{\frac{1}{2} + \frac{\eta}{8}} [(D_k u)_k(t_0, \cdot)]_{\alpha_k^k}. \end{aligned} \quad (3.46)$$

Let us now deal with the contribution $\mathcal{B}_1^{\xi', \tilde{\xi}'}(t_0, \mathbf{x}')$ in (3.42). Observe from this definition that this term is non zero if and only if $k < n$. Write then

$$\begin{aligned} |\mathcal{B}_1^{\xi', \tilde{\xi}'}(t_0, \mathbf{x}')| &\leq \int_{\mathbb{R}^{nd}} |D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y})| |u(t_0, \mathbf{y}) - u(t_0, \mathbf{y}_{1:k}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k+1:n})| d\mathbf{y} \\ &\quad + \int_{\mathbb{R}^{nd}} |D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y})| |u(t_0, \mathbf{y}) - u(t_0, \mathbf{y}_{1:k}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k+1:n})| d\mathbf{y} \\ &\leq C \|\mathbf{D}u\|_{\infty} \left(\int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(t_0 - t)^{k - \frac{1}{2}}} \hat{p}_{C^{-1}}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) |(\boldsymbol{\theta}_{t_0, t}(\mathbf{x}') - \mathbf{y})_{k+1:n}| \right. \\ &\quad \left. + \int_{\mathbb{R}^{nd}} \frac{d\mathbf{y}}{(t_0 - t)^{k - \frac{1}{2}}} \hat{p}_{C^{-1}}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) (|(\mathbf{m}_{t_0, t}^{\xi'}(\mathbf{x}') - \mathbf{y})_{k+1:n}| + |(\mathbf{m}_{t_0, t}^{\xi'}(\mathbf{x}') - \boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k+1:n}|) \right). \end{aligned} \quad (3.47)$$

To deal with the last contribution in the r.h.s., we will need some auxilliary lemmas already used in [CDrRH18] for Schauder estimates. Namely, analogously to Lemmas 1 and 3 therein, we have the following result.

Lemma 12. *There exists $\vartheta = \vartheta((\mathbf{A})) \in (0, 1)$ s.t. for all $j \in \llbracket k, n \rrbracket$:*

$$|(\mathbf{m}_{t_0, t}^{\xi'}(\mathbf{x}') - \boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_j| \leq c_0^{\vartheta} |(\mathbf{x} - \mathbf{x}')_i|^{\frac{2j-1}{2i-1}}. \quad (3.48)$$

In particular, recalling that $t_0 = t + c_0 |(\mathbf{x} - \mathbf{x}')_i|^{2/(2i-1)}$ with $|(\mathbf{x} - \mathbf{x}')_i| \leq 1$, we obtain

$$\begin{aligned} \frac{\|\mathbf{D}u\|_{\infty}}{(t_0 - t)^{k - \frac{1}{2}}} |(\mathbf{m}_{t_0, t}^{\xi'}(\mathbf{x}') - \boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k+1:n}| &\leq \frac{\|\mathbf{D}u\|_{\infty}}{c_0^{k - \frac{1}{2}} |(\mathbf{x} - \mathbf{x}')_i|^{\frac{2k-1}{2i-1}}} c_0^{\vartheta} |(\mathbf{x} - \mathbf{x}')_i|^{\frac{2(k+1)-1}{2i-1}} \\ &\leq \|\mathbf{D}u\|_{\infty} c_0^{\vartheta - (k - \frac{1}{2})} |(\mathbf{x} - \mathbf{x}')_i|^{\frac{2}{2i-1}} \\ &\leq \|\mathbf{D}u\|_{\infty} c_0^{\vartheta - (k - \frac{1}{2})} |(\mathbf{x} - \mathbf{x}')_i|^{\alpha_i^k}. \end{aligned}$$

Plugging the above control in (3.47) we obtain

$$|\mathcal{B}_1^{\xi', \tilde{\xi}'}(t_0, \mathbf{x}')| \leq C \|\mathbf{D}u\|_{\infty} \left\{ 2(t_0 - t) + c_0^{\vartheta - (k - \frac{1}{2})} |(\mathbf{x} - \mathbf{x}')_i|^{\alpha_i^k} \right\}. \quad (3.49)$$

Let us eventually control the term $\mathcal{B}_3^{\xi', \tilde{\xi}'}(t_0, \mathbf{x}')$ defined in (3.44) which we rewrite in the following way:

$$\begin{aligned}
\mathcal{B}_3^{\xi', \tilde{\xi}'}(t_0, \mathbf{x}') &= \left\{ \int_{\mathbb{R}^{nd}} d\mathbf{y} D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \left\langle [D_k u(t_0, \mathbf{y}_{1:k-1}, (\boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_{k:n}) - D_k u(t_0, \boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))], \right. \right. \\
&\quad \left. \left. (\mathbf{y} - \boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_k \right\rangle \right. \\
&\quad - \int_{\mathbb{R}^{nd}} d\mathbf{y} D_{\mathbf{x}_k} \tilde{p}^{\xi}(t, t_0, \mathbf{x}', \mathbf{y}) \left\langle [D_k u(t_0, \mathbf{y}_{1:k-1}, (\mathbf{m}_{t_0, t}^{\xi}(\mathbf{x}'))_k, \boldsymbol{\theta}_{t_0, t}(\mathbf{x}')_{k+1:n}) \right. \\
&\quad \left. - D_k u(t_0, (\mathbf{m}_{t_0, t}^{\xi}(\mathbf{x}'))_{1:k}, \boldsymbol{\theta}_{t_0, t}(\mathbf{x}')_{k+1:n})] (\mathbf{y} - \mathbf{m}_{t_0, t}^{\xi}(\mathbf{x}'))_k \right\rangle \left. \right\} \\
&+ \left\{ \int_{\mathbb{R}^{nd}} d\mathbf{y} D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \right. \\
&\quad \times \left\langle [D_k u(t_0, \boldsymbol{\theta}_{t_0, t}(\mathbf{x}')) - D_k u(t_0, (\mathbf{m}_{t_0, t}^{\xi}(\mathbf{x}'))_{1:k}, \boldsymbol{\theta}_{t_0, t}(\mathbf{x}')_{k+1:n}), (\mathbf{y} - \boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_k] \right\rangle \left. \right\} \\
&- \left\{ \int_{\mathbb{R}^{nd}} d\mathbf{y} D_{\mathbf{x}_k} \tilde{p}^{\xi}(t, t_0, \mathbf{x}', \mathbf{y}) \left\langle D_k u(t_0, (\mathbf{m}_{t_0, t}^{\xi}(\mathbf{x}'))_{1:k}, \boldsymbol{\theta}_{t_0, t}(\mathbf{x}')_{k+1:n}), (\mathbf{y} - \mathbf{m}_{t_0, t}^{\xi}(\mathbf{x}'))_k \right\rangle \right. \\
&\quad \left. - \int_{\mathbb{R}^{nd}} d\mathbf{y} D_{\mathbf{x}_k} \tilde{p}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \left\langle D_k u(t_0, (\mathbf{m}_{t_0, t}^{\xi}(\mathbf{x}'))_{1:k}, \boldsymbol{\theta}_{t_0, t}(\mathbf{x}')_{k+1:n}), (\mathbf{y} - \boldsymbol{\theta}_{t_0, t}(\mathbf{x}'))_k \right\rangle \right\},
\end{aligned}$$

where, thanks to Proposition 6 the last contribution is actually 0. For the first and second contributions in the above r.h.s. we have, thanks to the Hölder regularity of $\mathbf{x}_{1:k} \mapsto D_k u(\cdot, \mathbf{x}_{1:k}, \cdot)$, Proposition 5 and Lemma 12:

$$\begin{aligned}
\mathcal{B}_3^{\xi', \tilde{\xi}'}(t_0, \mathbf{x}') &\leq C \int_{\mathbb{R}^{nd}} d\mathbf{y} \left\{ \tilde{p}_{C^{-1}}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) + \tilde{p}_{C^{-1}}^{\xi}(t, t_0, \mathbf{x}', \mathbf{y}) \right\} \sum_{j=1}^{k-1} [(D_k u)_j(t_0, \cdot)]_{\alpha_j^k} (t_0 - t)^{\alpha_j^k (j - \frac{1}{2})} \\
&\quad + C \int_{\mathbb{R}^{nd}} d\mathbf{y} \tilde{p}_{C^{-1}}^{\xi'}(t, t_0, \mathbf{x}', \mathbf{y}) \sum_{j=1}^k [(D_k u)_j(t_0, \cdot)]_{\alpha_j^k} (t_0 - t)^{\alpha_j^k (j - \frac{1}{2})} \\
&\leq C \left(\sum_{j=1}^k [(D_k u)_j(t_0, \cdot)]_{\alpha_j^k} \right) (t_0 - t)^{\frac{1}{2} + \frac{\eta}{8}}, \tag{3.50}
\end{aligned}$$

with again the notation of (1.3). Thus, plugging estimates (3.45), (3.46), (3.49) and (3.50) in (3.41) we deduce that

$$\begin{aligned}
&|D_{\mathbf{x}_k} \tilde{P}_{t_0, t}^{\xi'} u(t_0, \mathbf{x}') - D_{\mathbf{x}_k} \tilde{P}_{t_0, t}^{\xi} u(t_0, \mathbf{x}')| \\
&\leq C (t_0 - t)^{\frac{1}{2} + \frac{\eta}{8}} [(D_k u)_k(t_0, \cdot)]_{\alpha_k^k} + C \|\mathbf{D}u\|_{\infty} \left\{ (t_0 - t) + c_0^{\vartheta - (k - \frac{1}{2})} |(\mathbf{x} - \mathbf{x}')_i|^{\alpha_i^k} \right\} \\
&\quad + C \left(\sum_{j=1}^k [(D_k u)_j(t_0, \cdot)]_{\alpha_j^k} \right) (t_0 - t)^{\frac{1}{2} + \frac{\eta}{8}} \\
&\leq C (t_0 - t)^{\frac{1}{2} + \frac{\eta}{8}} \|(D_k^2 u)_k(t_0, \cdot)\|_{B_{\infty, \infty}^{\alpha_k^k - 1}} + C \|\mathbf{D}u\|_{\infty} \left\{ (t_0 - t) + c_0^{\vartheta - (k - \frac{1}{2})} |(\mathbf{x} - \mathbf{x}')_i|^{\alpha_i^k} \right\} \\
&\quad + C \left(\sum_{j=1}^k \|(D_j D_k u)_j(t_0, \cdot)\|_{B_{\infty, \infty}^{\alpha_k^k - 1}} \right) (t_0 - t)^{\frac{1}{2} + \frac{\eta}{8}}. \tag{3.51}
\end{aligned}$$

The last inequality is a consequence of Proposition 14, stated in Appendix C.2, which allows to dominate the Hölder norms $\left([(D_k u)_j(t_0, \cdot)]_{\alpha_j^k} \right)_{2 \leq j \leq k \leq n}$ in terms of the Besov norms $\left(\|(D_j D_k u)_j(t_0, \cdot)\|_{B_{\infty, \infty}^{\alpha_j^k - 1}} \right)_{2 \leq j \leq k \leq n}$, with the notations of (3.29). This point is actually crucial to complete our circular argument.

From the definition of $t_0 = t + c_0 |(\mathbf{x} - \mathbf{x}')_i|^{2/(2i-1)}$ and $\alpha_i^k = (1 + \eta/4)/(2i - 1)$, recalling that $t_0 - t$ is small as well (i.e. $t_0 - t \leq C(t_0 - t)^{1/2 + \eta/8}$), we obtain from (3.51):

$$\begin{aligned}
& |D_{\mathbf{x}_k} \tilde{P}_{t_0, t}^{\xi'} u(t_0, \mathbf{x}') - D_{\mathbf{x}_k} \tilde{P}_{t_0, t}^{\xi} u(t_0, \mathbf{x}')| \\
& \leq C \left\{ c_0^{\frac{1}{2} + \frac{\eta}{8}} \left(\sum_{j=1}^k \|(D_j D_k u)_j(t_0, \cdot)\|_{B_{\infty, \infty}^{\alpha_j^k - 1}} \right) + (c_0^{\vartheta - (k - \frac{1}{2})} + c_0^{\frac{1}{2} + \frac{\eta}{8}}) \|\mathbf{D}u\|_{\infty} \right\} |(\mathbf{x} - \mathbf{x}')_i|^{\alpha_i^k}.
\end{aligned}$$

Thus, from (3.25), there exists $\tilde{\delta} := \tilde{\delta}(\mathbf{A}) \in (0, 1)$ such that

$$\begin{aligned}
& \left| (h_v \star D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})(z))|_{\partial S_i} \right| \\
& \leq v^{\frac{\alpha_i^k - 1}{2}} C (c_0^{-\tilde{\delta} - 1} \|\mathbf{D}u\|_{\infty} + c_0^{\tilde{\delta}} \max_{2 \leq \ell \leq m \leq n} \|(D_{\ell} D_m u)_{\ell}(t_0, \cdot)\|_{B_{\infty, \infty}^{\alpha_{\ell}^m - 1}}). \tag{3.52}
\end{aligned}$$

Conclusion: control of (3.22). Plugging (3.30), (3.40) and (3.52) into (3.22), (3.20), we eventually derive that for some positive $\tilde{\delta} := \tilde{\delta}(\mathbf{A}) > 0$:

$$\begin{aligned}
& \|D_{\mathbf{x}_i} D_{\mathbf{x}_k} u(t, \mathbf{x}_{1:i-1}, \cdot, \mathbf{x}_{i+1:n})\|_{B_{\infty, \infty}^{\alpha_i^k - 1}} \tag{3.53} \\
& \leq C \left(c_0^{-\tilde{\delta} - 1} \|\mathbf{D}u\|_{\infty} + \|\mathbf{D}D_{\mathbf{x}_1} u\|_{\infty} + (c_0^{\tilde{\delta}} + T^{\tilde{\delta}}) \sup_{s \in [0, T], 2 \leq \ell \leq m \leq n} \|(D_{\ell} D_m u)_{\ell}(s, \cdot)\|_{B_{\infty, \infty}^{\alpha_{\ell}^m - 1}} \right).
\end{aligned}$$

The main point to close our circular argument consists then in taking the supremums over w.r.t. $\mathbf{x}_{1:i-1}, \mathbf{x}_{i+1:n}, i, k$ and $t \in [0, T]$ on the l.h.s. and to tune the constant c_0 and the terminal time T in order to obtain $C(c_0^{\tilde{\delta}} + T^{\tilde{\delta}}) \leq 1/2$. We then derive that for all $2 \leq i \leq k \leq n$:

$$\sup_{t \in [0, T]} \|(D_i D_k u)_i(t, \cdot)\|_{B_{\infty, \infty}^{\alpha_i^k - 1}} \leq C \{ \|\mathbf{D}u\|_{\infty} + \|\mathbf{D}D_{\mathbf{x}_1} u\|_{\infty} \}, \tag{3.54}$$

which concludes the proof of Lemma 10.

3.3 Proof of Lemma 11

We follow the proof of Lemma 9, concentrating on the case $\ell \leq k - 1$, the specific case $\ell = k$ could be treated similarly considering the slightly different cancellation terms already discussed in Lemma 9. The quantity to estimate is now:

$$\left[\Psi_{\ell, k, m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x})}(\cdot) - \Psi_{\ell, k, m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x}')}(\cdot) \right]. \tag{3.55}$$

Splitting the thermic part of the Besov norm as in (3.1), we obtain the same kind estimate for the non-singular in time part. Indeed, we point out that the difference (3.55) does not involve $D_{\mathbf{x}_1}$, therefore we gain a singularity of order $1/2$ w.r.t. to equation (3.2). On the other hand, the difference of the derivatives of the frozen densities w.r.t. \mathbf{x}_k can be handled with (3.33). Choosing $\rho_{\ell, m} = 2\ell - 1$ as in the proof of Lemma 9, and recalling that $\alpha_i^k(i - 1/2) = (1 + \eta/4)/2$ (see (3.13)), it is plain to check that:

$$\begin{aligned}
& \text{Upper} \left[\Psi_{\ell, k, m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x})}(\cdot) - \Psi_{\ell, k, m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x}')}(\cdot) \right] \\
& \leq \frac{\tilde{C} |(\mathbf{x} - \mathbf{x}')_i|^{\alpha_i^k}}{(s - t)^{1 + \alpha_i^k(i - \frac{1}{2}) - \gamma_{\ell}^m}} \hat{q}_{c \setminus \ell}(t, s, \mathbf{x}, (\mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n})), \tag{3.56}
\end{aligned}$$

where from (\mathbf{H}_{η}) , $1 + \alpha_i^k(i - 1/2) - \gamma_{\ell}^m < 1$. Turning to the singular in time contribution of the thermic part of the Besov norm of (3.55) we decompose with the notations of Lemma 9 (see e.g. (3.7), (3.8) which exhibit an additional spatial derivative):

$$\begin{aligned}
& \text{Lower} \left[\Psi_{\ell, k, m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x})}(\cdot) - \Psi_{\ell, k, m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x}')}(\cdot) \right] \\
& =: \int_0^{(s-t)^{\rho_{\ell, m}}} dw w^{\frac{\alpha_{\ell}^m - 1}{2}} \int_{\mathbb{R}^d} d\tilde{z} |(\mathcal{F}_{1, \ell, k, m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x}, \mathbf{x}')} + \mathcal{F}_{2, \ell, k, m}^{(s, \mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n}), (t, \mathbf{x}, \mathbf{x}')})(w, \tilde{z})|,
\end{aligned}$$

where $\ell < k \leq m$ and:

$$\begin{aligned} & \mathcal{J}_{1,\ell,k,m}^{(s,\mathbf{y}_{1:\ell-1},\mathbf{y}_{\ell+1:n}), (t,\mathbf{x},\mathbf{x}')} (w, \tilde{z}) \\ &= \int_{\mathbb{R}^d} \partial_w h_w(\tilde{z} - \mathbf{y}_\ell) (D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}', \mathbf{y})) \\ & \quad \langle \mathbf{F}_\ell(s, \mathbf{y}_{1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})) - \mathbf{F}_\ell(s, \mathbf{y}_{1:\ell-1}, \tilde{z}, \mathbf{y}_{\ell+1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})), (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_m \rangle d\mathbf{y}_\ell, \end{aligned} \quad (3.57)$$

with a slight abuse of notation when $\ell = k - 1$ and

$$\begin{aligned} & \mathcal{J}_{2,\ell,k,m}^{(s,\mathbf{y}_{1:\ell-1},\mathbf{y}_{\ell+1:n}), (t,\mathbf{x})} (w, \tilde{z}) \\ &= \int_{\mathbb{R}^d} \partial_w h_w(\tilde{z} - \mathbf{y}_\ell) \left[\left(D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}_{1:\ell-1}, \tilde{z}, \mathbf{y}_{\ell+1:n}) \right) \right. \\ & \quad \left. - \left(D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}', \mathbf{y}) - D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}', \mathbf{y}_{1:\ell-1}, \tilde{z}, \mathbf{y}_{\ell+1:n}) \right) \right] \\ & \quad \langle \mathbf{F}_\ell(s, \mathbf{y}_{1:\ell-1}, \tilde{z}, \mathbf{y}_{\ell+1:k-1}, \boldsymbol{\theta}_{s,t}^{k:n}(\boldsymbol{\xi})) - \mathbf{F}_\ell(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) - D_{\ell-1} \mathbf{F}_\ell(s, \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi})) (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_{\ell-1}, \\ & \quad (\mathbf{y} - \boldsymbol{\theta}_{s,t}(\boldsymbol{\xi}))_m \rangle d\mathbf{y}_\ell. \end{aligned} \quad (3.58)$$

Note now that when proceeding first as in (3.31), (3.32) and then control the resulting difference as in (3.8) we get, thanks to (3.33), that:

$$\begin{aligned} & \left| \left(D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}) - D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}, \mathbf{y}_{1:\ell-1}, \tilde{z}, \mathbf{y}_{\ell+1:n}) \right) \right. \\ & \quad \left. - \left(D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}', \mathbf{y}) - D_{\mathbf{x}_k} \tilde{p}^\xi(t, s, \mathbf{x}', \mathbf{y}_{1:\ell-1}, \tilde{z}, \mathbf{y}_{\ell+1:n}) \right) \right| \\ & \leq \frac{\tilde{C}}{(s-t)^{\ell-\frac{1}{2}+i-\frac{1}{2}+k-\frac{1}{2}}} \int_0^1 d\lambda \tilde{p}_{C-1}^\xi(t, s, \mathbf{x}, \mathbf{y}_{1:\ell-1}, \tilde{z} + \lambda(\mathbf{y}_\ell - z), \mathbf{y}_{\ell+1:n}) |(\mathbf{x}' - \mathbf{x})_i| |\tilde{z} - \mathbf{y}_\ell| \\ & \leq \frac{\tilde{C}}{(s-t)^{(\ell-\frac{1}{2})+\alpha_i^k(i-\frac{1}{2})+(k-\frac{1}{2})}} \int_0^1 d\lambda \tilde{p}_{C-1}^\xi(t, s, \mathbf{x}, \mathbf{y}_{1:\ell-1}, \tilde{z} + \lambda(\mathbf{y}_\ell - z), \mathbf{y}_{\ell+1:n}) |(\mathbf{x}' - \mathbf{x})_i|^{\alpha_i^k} |\tilde{z} - \mathbf{y}_\ell|, \end{aligned}$$

using the fact we are in the diagonal regime in the last inequality. With this control at hand, together with estimate (3.33), to handle the contributions $(\mathcal{J}_{1,\ell,k,m}^{(s,\mathbf{y}_{1:\ell-1},\mathbf{y}_{\ell+1:n}), (t,\mathbf{x})} + \mathcal{J}_{2,\ell,k,m}^{(s,\mathbf{y}_{1:\ell-1},\mathbf{y}_{\ell+1:n}), (t,\mathbf{x})})(w, \tilde{z})$, we can mimic the proof of the estimation of the contributions $(\mathcal{J}_{1,i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})} + \mathcal{J}_{2,i,(l,1),k}^{(s,\mathbf{y}_{1:i-1},\mathbf{y}_{i+1:n}), (t,\mathbf{x})})(v, z)$ done in Lemma 9 to obtain

$$\begin{aligned} & \mathbf{Lower} \left[\Psi_{\ell,k,m}^{(s,\mathbf{y}_{1:\ell-1},\mathbf{y}_{\ell+1:n}), (t,\mathbf{x})}(\cdot) - \Psi_{\ell,k,m}^{(s,\mathbf{y}_{1:\ell-1},\mathbf{y}_{\ell+1:n}), (t,\mathbf{x}')}(\cdot) \right] \\ & \leq \tilde{C} \hat{q}_{c\setminus\ell}(t, s, \mathbf{x}, (\mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n})) |(\mathbf{x}' - \mathbf{x})_i|^{\alpha_i^k} \int_0^{(s-t)^{\rho_{\ell,m}}} dw w^{\frac{\alpha_{\ell}^m - 1}{2}} \\ & \quad \left\{ \frac{1}{w^{1-\frac{\beta_\ell}{2}} (s-t)^{\alpha_i^k(i-\frac{1}{2})}} + \frac{w^{-\frac{1+\beta_\ell}{2}}}{(s-t)^{\ell-\frac{1}{2}+\alpha_i^k(i-\frac{1}{2})}} + \frac{w^{-\frac{1}{2}}}{(s-t)^{\ell-\frac{1}{2}+\alpha_i^k(i-\frac{1}{2})-(1+\eta)(\ell-3/2)}} \right\} \\ & \leq \tilde{C} \hat{q}_{c\setminus\ell}(t, s, \mathbf{x}, (\mathbf{y}_{1:\ell-1}, \mathbf{y}_{\ell+1:n})) (s-t)^{-1-(i-\frac{1}{2})\alpha_i^k+\gamma_\ell^m} |(\mathbf{x}' - \mathbf{x})_i|^{\alpha_i^k}, \end{aligned}$$

where $\gamma_\ell^m = 1/2 + \eta(\ell - 3/2)$. From (\mathbf{H}_η) and the very definition of α_i^k we hence have $1 - \alpha_i^k(i - 1/2) + \gamma_\ell^m < 1$ which, together with (3.56), concludes the proof. \square

A Sensitivity results for the mean: Proof of Lemma 12

In order to prove Lemma 12, we first need to establish some controls on the sensitivity of the flows, see Lemma 13 below. Those results are obtained under the sole assumption **(A)** and remain valid for the mollification

procedure of the coefficients considered in **(AM)**. We will then proceed to the final proof of Lemma 12 in Section A.2.

For our analysis, we now introduce the spatial homogeneous distance, which basically reflects the various scales of the system already seen e.g. in Proposition 5. Namely, for $(\mathbf{x}, \mathbf{x}') \in \mathbb{R}^{nd}$, we define:

$$\mathbf{d}(\mathbf{x}, \mathbf{x}') := \sum_{i=1}^n |(\mathbf{x} - \mathbf{x}')_i|^{\frac{1}{2i-1}}. \quad (\text{A.1})$$

The distance is homogeneous in the sense that, for any $\lambda > 0$, $\mathbf{d}(\lambda^{-1/2}\mathbb{T}_\lambda \mathbf{x}, \lambda^{-1/2}\mathbb{T}_\lambda \mathbf{x}') = \lambda^{1/2}\mathbf{d}(\mathbf{x}, \mathbf{x}')$.

A.1 A first sensitivity result for the flow

Lemma 13 (Control of the flows). *Under **(A)**, there exists $C := C((\mathbf{A}), T)$ s.t. for all spatial points $(\mathbf{x}, \mathbf{x}') \in (\mathbb{R}^{nd})^2$, $\mathbf{d}(\mathbf{x}, \mathbf{x}') \leq 1$, $0 \leq t < s \leq T \leq 1$ and $i \in \llbracket 1, n \rrbracket$:*

$$|(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_i| \leq C \left((s-t)^{i-\frac{1}{2}} + \mathbf{d}^{2i-1}(\mathbf{x} - \mathbf{x}') \right).$$

The flow, $\boldsymbol{\theta}_{s,t}$ is, somehow, locally ‘‘almost’’ Lipschitz continuous in space w.r.t. the homogeneous distance \mathbf{d} , up to a time additive term. This time contribution is a consequence of the non Lipschitz continuity of the drift \mathbf{F} . The analysis which was already done for \mathbf{F} Lipschitz continuous in Proposition 4.1 of [Men18], and Appendix A.1 in [CdRHM18] with different Hölder regularity of \mathbf{F} . Actually, as we consider a *smoother* drift than in [CdRHM18], the following lemma can be seen as a by-product of Lemma 12 therein. For the sake of completeness, we provide the corresponding, and more direct, analysis below.

Proof. The analysis mainly relies on Grönwall type arguments coupled with suitable mollification techniques, because \mathbf{F} is not Lipschitz continuous, and appropriate Young inequalities in order to make the intrinsic scales associated to the spatial variables appear. Let $\delta \in \mathbb{R}^n$ be the vector whose entries $\delta_i > 0$ correspond to the mollification parameter of the drift \mathbf{F}_i for $i \in \llbracket 2, n \rrbracket$. Namely, for all $v \in [0, T]$, $\mathbf{z} \in \mathbb{R}^{nd}$, $i \in \llbracket 2, n \rrbracket$, we define

$$\mathbf{F}_i^\delta(v, \mathbf{z}^{i-1:n}) := \mathbf{F}_i(v, \cdot) \star \rho_{\delta_i}(\mathbf{z}) = \int_{\mathbb{R}^d} \mathbf{F}_i(v, \mathbf{z}_{i-1}, \mathbf{z}_i - w, \mathbf{z}_{i+1}, \dots, \mathbf{z}_n) \rho_{\delta_i}(w) dw, \quad (\text{A.2})$$

with $\rho_{\delta_i}(w) := (1/\delta_i^d)\rho(w/\delta_i)$ where $\rho : \mathbb{R}^d \rightarrow \mathbb{R}_+$ is a usual mollifier, namely ρ has compact support and $\int_{\mathbb{R}^d} \rho(\mathbf{z}) d\mathbf{z} = 1$. Eventually, we write $\mathbf{F}^\delta(v, \mathbf{z}) := (\mathbf{F}_1(v, \mathbf{z}), \mathbf{F}_2^\delta(v, \mathbf{z}), \dots, \mathbf{F}_n^\delta(v, \mathbf{z}))$. With a slight abuse of notation in the previous definitions, since the first component \mathbf{F}_1 is not mollified. The sublinearity of \mathbf{F}_1 is actually enough to obtain the desired control.

To be at the *good* current time scale for the contributions associated with the mollification, we pick δ_i in order to have $C := C((\mathbf{A}), T) > 0$ s.t. for all $\mathbf{z} \in \mathbb{R}^{nd}$, $u \in [t, s]$:

$$\left| (s-t)^{\frac{1}{2}} \mathbb{T}_{s-t}^{-1} \left(\mathbf{F}(u, \mathbf{z}) - \mathbf{F}^\delta(u, \mathbf{z}) \right) \right| \leq C(s-t)^{-1}. \quad (\text{A.3})$$

By the previous definition of \mathbf{F}^δ in (A.2), identity (A.3) is equivalent to:

$$\sum_{i=2}^n (s-t)^{\frac{1}{2}-i} \delta_i^{\frac{2i-2}{2i-1}} \leq C(s-t)^{-1}. \quad (\text{A.4})$$

Hence, we choose from now on, for all $i \in \llbracket 2, n \rrbracket$:

$$\delta_i = (s-t)^{(i-\frac{3}{2})\frac{2i-1}{2i-2}}. \quad (\text{A.5})$$

Next, let us control the last components of the flow. By the definition of $\boldsymbol{\theta}_{s,t}$ in (2.4), we get:

$$\begin{aligned} & |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_n| \\ & \leq |(\mathbf{x} - \mathbf{x}')_n| + \int_t^s \left(|\mathbf{F}_n^\delta(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) - \mathbf{F}_n^\delta(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}'))| \right. \\ & \quad \left. + |\mathbf{F}_n^\delta(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) - \mathbf{F}_n(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}))| + |\mathbf{F}_n^\delta(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}')) - \mathbf{F}_n(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}'))| \right) dv \\ & \leq |(\mathbf{x} - \mathbf{x}')_n| + C \int_t^s \left(|(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-1}| + \delta_n^{-1+\frac{2n-2}{2n-1}} |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_n| \right) dv \\ & \quad + (s-t) \delta_n^{\frac{2n-2}{2n-1}}, \end{aligned}$$

observing for the last inequality that since $\beta_n > (2n-2)/(2n-1)$ and δ_n is meant to be small, $\delta_n^{\beta_n} \leq \delta_n^{(2n-2)/(2n-1)}$.

Hence by Grönwall's Lemma, we get:

$$\begin{aligned} & |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_n| \\ & \leq C \exp\left(C(s-t)\delta_n^{-1+\frac{2n-2}{2n-1}}\right) \left(|(\mathbf{x} - \mathbf{x}')_n| + (s-t)\delta_n^{\frac{2n-2}{2n-1}} + \int_t^s |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-1}| dv\right) \\ & \leq C \exp\left(C(s-t)^{\frac{1}{2}}\right) \left(|(\mathbf{x} - \mathbf{x}')_n| + (s-t)^{n-\frac{1}{2}} + \int_t^s |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-1}| dv\right), \end{aligned} \quad (\text{A.6})$$

using (A.5) for the last inequality. For the $(n-1)^{\text{th}}$ component, the situation is quite different in the sense that we have to handle the non-Lipschitz continuity of \mathbf{F}_{n-1}^δ in its n^{th} variable. Write:

$$\begin{aligned} & |(\boldsymbol{\theta}_{s,t}(\mathbf{x}) - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_{n-1}| \\ & \leq C \exp\left(C(s-t)\delta_{n-1}^{-1+\frac{2(n-1)-2}{2(n-1)-1}}\right) \left(|(\mathbf{x} - \mathbf{x}')_{n-1}| + (s-t)\delta_{n-1}^{\frac{2(n-1)-2}{2(n-1)-1}}\right. \\ & \quad \left. + \int_t^s \left\{ |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-2}| + |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_n|^{\beta_n} \right\} dv\right) \\ & \leq C \exp\left(C(s-t)^{\frac{1}{2}}\right) \left(|(\mathbf{x} - \mathbf{x}')_{n-1}| + (s-t)^{n-\frac{3}{2}} + \int_t^s \left\{ |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-2}| \right. \right. \\ & \quad \left. \left. + |(\mathbf{x} - \mathbf{x}')_n|^{\beta_n} + (v-t)^{\beta_n(n-\frac{1}{2})} + \left(\int_t^v |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}| dw \right)^{\beta_n} \right\} dv\right). \\ & \leq C \exp\left(C(s-t)^{\frac{1}{2}}\right) \left(|(\mathbf{x} - \mathbf{x}')_{n-1}| + (s-t)^{n-\frac{3}{2}} + \int_t^s \left\{ |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-2}| \right. \right. \\ & \quad \left. \left. + |(\mathbf{x} - \mathbf{x}')_n|^{\frac{2n-3}{2n-1}} + \left(\int_t^v |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}| dw \right)^{\beta_n} \right\} dv\right), \end{aligned} \quad (\text{A.7})$$

from our choice of δ_{n-1} in (A.5) for the second inequality. We also exploited for the last inequality that, since under (\mathbf{T}_β) , $\beta_n > (2n-2)/(2n-1)$, $\beta_n(n-\frac{1}{2}) > n-1 > n-3/2$ and $0 \leq t < s \leq T$ where T is *small*, then $(v-t)^{\beta_n(n-\frac{1}{2})} \leq (v-t)^{n-3/2}$. Also, since $\mathbf{d}(\mathbf{x}, \mathbf{x}') \leq 1$, the same arguments yield $|(\mathbf{x} - \mathbf{x}')_n|^{\beta_n} \leq |(\mathbf{x} - \mathbf{x}')_n|^{(2n-2)/(2n-1)} \leq |(\mathbf{x} - \mathbf{x}')_n|^{(2n-3)/(2n-1)}$.

From (A.7), which still holds true replacing s by any $\tilde{s} \in [t, s]$, we deduce that taking the supremum over $\tilde{s} \in [t, s]$:

$$\begin{aligned} & \sup_{\tilde{s} \in [t, s]} |(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_{n-1}| \\ & \leq C \exp\left(C(s-t)^{\frac{1}{2}}\right) \left(|(\mathbf{x} - \mathbf{x}')_{n-1}| + (s-t)^{n-\frac{3}{2}} + \int_t^s \left\{ |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-2}| \right. \right. \\ & \quad \left. \left. + |(\mathbf{x} - \mathbf{x}')_n|^{\frac{2n-3}{2n-1}} + \left(\int_t^v |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}| dw \right)^{\beta_n} \right\} dv\right). \end{aligned}$$

Taking then the supremum in $w \in [t, s]$ in the above integral, we obtain:

$$\begin{aligned} & \sup_{\tilde{s} \in [t, s]} |(\boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s},t}(\mathbf{x}'))_{n-1}| \\ & \leq C \exp\left(C(s-t)^{\frac{1}{2}}\right) \left(|(\mathbf{x} - \mathbf{x}')_{n-1}| + (s-t)^{n-\frac{3}{2}} + \int_t^s |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{n-2}| dv \right. \\ & \quad \left. + |(\mathbf{x} - \mathbf{x}')_n|^{\frac{2n-3}{2n-1}} + \sup_{w \in [t, s]} |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}|^{\beta_n} (s-t)^{\beta_n+1}\right). \end{aligned} \quad (\text{A.8})$$

From Young's inequality we now derive:

$$\begin{aligned} \sup_{w \in [t, s]} |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}|^{\beta_n} (s-t)^{\beta_n+1} & \leq \frac{1}{2} \sup_{w \in [t, s]} |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}| (s-t) + C(s-t)^{\frac{1}{1-\beta_n}} \\ & \leq \frac{1}{2} \sup_{w \in [t, s]} |(\boldsymbol{\theta}_{w,t}(\mathbf{x}) - \boldsymbol{\theta}_{w,t}(\mathbf{x}'))_{n-1}| + C(s-t)^{n-\frac{3}{2}}, \end{aligned}$$

recalling for the last inequality that $s - t \leq 1$, and since $\beta_n > (2n - 2)/(2n - 1)$, $1 - \beta_n < 1/(2n - 1)$, we also have $(s - t)^{1/(1-\beta_n)} < (s - t)^{2n-1} < (s - t)^{n-3/2}$. Plugging the above control into (A.8), we obtain:

$$\begin{aligned} & \frac{1}{2} \sup_{\tilde{s} \in [t, s]} |(\boldsymbol{\theta}_{\tilde{s}, t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s}, t}(\mathbf{x}'))_{n-1}| \\ & \leq C \exp(C(s - t)^{\frac{1}{2}}) \left(|(\mathbf{x} - \mathbf{x}')_{n-1}| + (s - t)^{n-\frac{3}{2}} + \int_t^s |(\boldsymbol{\theta}_{v, t}(\mathbf{x}) - \boldsymbol{\theta}_{v, t}(\mathbf{x}'))_{n-2}| dv \right. \\ & \quad \left. + |(\mathbf{x} - \mathbf{x}')_n|^{\frac{2n-3}{2n-1}} \right). \end{aligned} \quad (\text{A.9})$$

We explicitly see from (A.9) that each entry of the difference of the starting points appears at its intrinsic scale for the homogeneous distance \mathbf{d} introduced in (A.1).

Plugging the above inequality into (A.6) we derive:

$$\begin{aligned} |(\boldsymbol{\theta}_{s, t}(\mathbf{x}) - \boldsymbol{\theta}_{s, t}(\mathbf{x}'))_n| & \leq C \exp\left(C(s - t)^{\frac{1}{2}}\right) \left(|(\mathbf{x} - \mathbf{x}')_n| + (s - t)^{n-\frac{1}{2}} + |(\mathbf{x} - \mathbf{x}')_{n-1}|(s - t) \right. \\ & \quad \left. + |(\mathbf{x} - \mathbf{x}')_n|^{\frac{2n-3}{2n-1}}(s - t) + \int_t^s \int_t^v |(\boldsymbol{\theta}_{w, t}(\mathbf{x}) - \boldsymbol{\theta}_{w, t}(\mathbf{x}'))_{n-2}| dw dv \right) \\ & \leq C \exp\left(C(s - t)^{\frac{1}{2}}\right) \left(|(\mathbf{x} - \mathbf{x}')_n| + (s - t)^{n-\frac{1}{2}} + |(\mathbf{x} - \mathbf{x}')_{n-1}|^{\frac{2n-1}{2n-3}} \right. \\ & \quad \left. + \int_t^s \int_t^v |(\boldsymbol{\theta}_{w, t}(\mathbf{x}) - \boldsymbol{\theta}_{w, t}(\mathbf{x}'))_{n-2}| dw dv \right), \end{aligned}$$

using again the Young inequalities $|(\mathbf{x} - \mathbf{x}')_n|^{(2n-3)/(2n-1)}(s - t) \leq C(|(\mathbf{x} - \mathbf{x}')_n| + (s - t)^{n-1/2})$ and $|(\mathbf{x} - \mathbf{x}')_{n-1}|(s - t) \leq C(|(\mathbf{x} - \mathbf{x}')_{n-1}|^{(2n-1)/(2n-3)} + (s - t)^{n-1/2})$ for the last inequality. Iterating the procedure, we get:

$$\begin{aligned} & |(\boldsymbol{\theta}_{s, t}(\mathbf{x}) - \boldsymbol{\theta}_{s, t}(\mathbf{x}'))_n| \quad (\text{A.10}) \\ & \leq C \left((s - t)^{n-\frac{1}{2}} + \sum_{j=2}^n |(\mathbf{x} - \mathbf{x}')_j|^{\frac{2n-1}{2j-1}} + \int_t^{v_n=s} dv_{n-1} \dots \int_t^{v_2} dv_1 |(\boldsymbol{\theta}_{v_1, t}(\mathbf{x}') - \boldsymbol{\theta}_{v_1, t}(\mathbf{x}))_1| \right). \end{aligned}$$

Analogously, for $i \in [2, n]$, we obtain:

$$\begin{aligned} & |(\boldsymbol{\theta}_{s, t}(\mathbf{x}) - \boldsymbol{\theta}_{s, t}(\mathbf{x}'))_i| \quad (\text{A.11}) \\ & \leq C \left((s - t)^{i-\frac{1}{2}} + \sum_{j=2}^n |(\mathbf{x} - \mathbf{x}')_j|^{\frac{2i-1}{2j-1}} + \int_t^{v_i=s} dv_{i-1} \dots \int_t^{v_2} dv_1 |(\boldsymbol{\theta}_{v_1, t}(\mathbf{x}') - \boldsymbol{\theta}_{v_1, t}(\mathbf{x}))_1| \right). \end{aligned}$$

Remark 5. Observe that equations (A.10) and (A.11) are available for any fixed time $s \in [t, T]$.

The first term, i.e. for $i = 1$ is controlled slightly differently. In other words, for all $\tilde{s} \in [t, s]$, write:

$$|(\boldsymbol{\theta}_{\tilde{s}, t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s}, t}(\mathbf{x}'))_1| \leq |(\mathbf{x} - \mathbf{x}')_1| + C \sum_{j=1}^n \int_t^{\tilde{s}} |(\boldsymbol{\theta}_{v, t}(\mathbf{x}) - \boldsymbol{\theta}_{v, t}(\mathbf{x}'))_j|^{\beta_j} dv,$$

which in turn implies from (A.11), Remark 5 and convexity inequalities:

$$\begin{aligned} \sup_{\tilde{s} \in [t, s]} |(\boldsymbol{\theta}_{\tilde{s}, t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s}, t}(\mathbf{x}'))_1| & \leq |(\mathbf{x} - \mathbf{x}')_1| + C \left((s - t) \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v, t}(\mathbf{x}) - \boldsymbol{\theta}_{v, t}(\mathbf{x}'))_1|^{\beta_1} \right. \\ & \quad \left. + \sum_{j=2}^n \int_t^s |(\boldsymbol{\theta}_{v, t}(\mathbf{x}) - \boldsymbol{\theta}_{v, t}(\mathbf{x}'))_j|^{\beta_j} dv \right) \\ & \leq |(\mathbf{x} - \mathbf{x}')_1| + \left((s - t) \left[\frac{1}{2} \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{\tilde{s}, t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s}, t}(\mathbf{x}'))_1| + C \right] \right. \\ & \quad \left. + C \sum_{j=2}^n (s - t) \left((s - t)^{j-\frac{1}{2}} + \sum_{k=2}^n |(\mathbf{x} - \mathbf{x}')_k|^{\frac{2j-1}{2k-1}} \right) \right. \\ & \quad \left. + (s - t)^{j-1} \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v, t}(\mathbf{x}) - \boldsymbol{\theta}_{v, t}(\mathbf{x}'))_1|^{\beta_j} \right). \end{aligned}$$

Recalling $\beta_j > (2j - 2)/(2j - 1)$ and $0 \leq t < s \leq T \leq 1$, $\mathbf{d}(\mathbf{x}, \mathbf{x}') \leq 1$, we get:

$$\begin{aligned} \frac{1}{2} \sup_{\tilde{s} \in [t, s]} |(\boldsymbol{\theta}_{\tilde{s}, t}(\mathbf{x}) - \boldsymbol{\theta}_{\tilde{s}, t}(\mathbf{x}'))_1| &\leq |(\mathbf{x} - \mathbf{x}')_1| + C \left((s - t) \right. \\ &\quad \left. + \sum_{j=2}^n (s - t) \left((s - t)^{\frac{1}{2}} + \sum_{k=2}^n |(\mathbf{x} - \mathbf{x}')_k|^{\frac{1}{2k-1}} \right) \right. \\ &\quad \left. + (s - t)^{(j-1)\beta_j} \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v, t}(\mathbf{x}) - \boldsymbol{\theta}_{v, t}(\mathbf{x}'))_1|^{\beta_j} \right). \end{aligned} \quad (\text{A.12})$$

Write now from the Young inequality:

$$C(s - t)^{1+(j-1)\beta_j} \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v, t}(\mathbf{x}) - \boldsymbol{\theta}_{v, t}(\mathbf{x}'))_1|^{\beta_j} \leq C(s - t) + \frac{1}{4} \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v, t}(\mathbf{x}) - \boldsymbol{\theta}_{v, t}(\mathbf{x}'))_1|.$$

We eventually derive from (A.12) that:

$$\sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v, t}(\mathbf{x}) - \boldsymbol{\theta}_{v, t}(\mathbf{x}'))_1| \leq C \left((s - t)^{\frac{1}{2}} + \mathbf{d}(\mathbf{x}, \mathbf{x}') \right),$$

which gives the statement for $i = 1$. Plugging now this inequality into (A.11), we get for all $i \in \llbracket 2, n \rrbracket$:

$$\begin{aligned} &|(\boldsymbol{\theta}_{s, t}(\mathbf{x}) - \boldsymbol{\theta}_{s, t}(\mathbf{x}'))_i| \\ &\leq C \left((s - t)^{i-\frac{1}{2}} + \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}') + (s - t)^{i-1} \sup_{v \in [t, s]} |(\boldsymbol{\theta}_{v, t}(\mathbf{x}) - \boldsymbol{\theta}_{v, t}(\mathbf{x}'))_1| \right) \\ &\leq C \left((s - t)^{i-\frac{1}{2}} + \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}') + (s - t)^{i-1} \left((s - t)^{\frac{1}{2}} + \mathbf{d}(\mathbf{x}, \mathbf{x}') \right) \right) \\ &\leq C \left((s - t)^{i-\frac{1}{2}} + \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}') \right), \end{aligned}$$

using again for the last identity the Young inequality to derive that $(s - t)^{i-1} \mathbf{d}(\mathbf{x}, \mathbf{x}') \leq C \left((s - t)^{i-1/2} + \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}') \right)$. The proof is complete. \square

A.2 Sensitivity results for the mean: final proof of Lemma 12

Again through the analysis, we assume w.l.o.g. that $\mathbf{d}(\mathbf{x}, \mathbf{x}') \leq 1$. The control is done with a distinction of two contributions to handle.

$$\mathbf{m}_{s, t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s, t}(\mathbf{x}') = [\mathbf{m}_{s, t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s, t}(\mathbf{x})] + [\boldsymbol{\theta}_{s, t}(\mathbf{x}) - \boldsymbol{\theta}_{s, t}(\mathbf{x}')]. \quad (\text{A.13})$$

By the proxy definition in (2.3), we deduce that the mean value of $\tilde{\mathbf{X}}_v^{m, \xi}$, $\mathbf{m}_{v, t}^{\xi}$ is s.t.

$$\mathbf{m}_{s, t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s, t}(\mathbf{x}) = \mathbf{x}' - \mathbf{x} + \int_t^s dv D\mathbf{F}(v, \boldsymbol{\theta}_{v, t}(\mathbf{x})) [\mathbf{m}_{v, t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{v, t}(\mathbf{x})]. \quad (\text{A.14})$$

The sub-triangular structure of $D\mathbf{F}$ yields that for all $i \in \llbracket 2, n \rrbracket$:

$$(\mathbf{m}_{s, t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s, t}(\mathbf{x}))_i = \mathbf{x}'_i - \mathbf{x}_i + \int_t^s dv D_{i-1} \mathbf{F}_i(v, \boldsymbol{\theta}_{v, t}(\mathbf{x})) [\mathbf{m}_{v, t}^{\mathbf{x}}(\mathbf{x}')_{i-1} - \boldsymbol{\theta}_{v, t}(\mathbf{x})_{i-1}].$$

Also, since $\mathbf{m}_{v, t}^{\mathbf{x}}(\mathbf{x}')_1 = \mathbf{x}'_1 + \int_t^s \mathbf{F}_1(v, \boldsymbol{\theta}_{v, t}(\mathbf{x})) dv$, so we obtain that $[\mathbf{m}_{v, t}^{\mathbf{x}}(\mathbf{x}')_1 - \boldsymbol{\theta}_{v, t}(\mathbf{x})_1] = \mathbf{x}'_1 - \mathbf{x}_1$, we then obtain by iteration that:

$$\begin{aligned} &(\mathbf{m}_{s, t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s, t}(\mathbf{x}))_i \\ &= \mathbf{x}'_i - \mathbf{x}_i + \sum_{k=2}^i \left[\int_t^{v_i=s} dv_{i-1} \dots \int_t^{v_k} dv_{k-1} \prod_{j=k}^i D_{j-1} \mathbf{F}_j(v_j, \boldsymbol{\theta}_{v_j, t}(\mathbf{x})) \right] [\mathbf{x}'_{k-1} - \mathbf{x}_{k-1}], \end{aligned}$$

with the convention that for $i = 1$, $\sum_{k=2}^i = 0$. From the above control, equation (A.13) and the dynamics of the flow, and because the starting points are the same, the contributions involving differences of the spatial

points $(\mathbf{x}' - \mathbf{x})$ or flows only appear in iterated time integrals, we obtain:

$$\begin{aligned}
& |(\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_i| \\
& \leq \left| \sum_{k=2}^i \left[\int_t^{v_i=s} dv_{i-1} \dots \int_t^{v_k} dv_{k-1} \prod_{j=k}^i D_{j-1} \mathbf{F}_j(v_j, \boldsymbol{\theta}_{v_j,t}(\mathbf{x})) \right] [\mathbf{x}'_{k-1} - \mathbf{x}_{k-1}] \right| \\
& \quad + \int_t^s |\mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x})) - \mathbf{F}_i(v, \boldsymbol{\theta}_{v,t}(\mathbf{x}'))| dv \\
& \leq C \left(\sum_{k=2}^{i-1} (s-t)^{i-k} |\mathbf{x}_k - \mathbf{x}'_k| \right. \\
& \quad \left. + \int_t^s \left(\sum_{j=i}^n |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_j|^{\beta_j} + |(\boldsymbol{\theta}_{v,t}(\mathbf{x}) - \boldsymbol{\theta}_{v,t}(\mathbf{x}'))_{i-1}| \right) dv \right).
\end{aligned}$$

We derive from the previous Lemma 13 (control of the flows) recalling again that $\beta_j > (2j-2)/(2j-1)$ and $\mathbf{d}(\mathbf{x}, \mathbf{x}') \leq 1, 0 \leq t < s \leq T \leq 1$:

$$\begin{aligned}
& |(\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{s,t}(\mathbf{x}'))_i| \\
& \leq C \left(\sum_{k=2}^{i-1} (s-t)^{i-k} |\mathbf{x}_k - \mathbf{x}'_k| + (s-t)^{\frac{2i-2}{2}+1} \right. \\
& \quad \left. + \mathbf{d}^{2i-2}(\mathbf{x}, \mathbf{x}') (s-t) + ((s-t)^{(i-1)-\frac{1}{2}} + \mathbf{d}^{2(i-1)-1}(\mathbf{x}, \mathbf{x}')) (s-t) \right).
\end{aligned}$$

In particular, for $s = t_0 = t + c_0 \mathbf{d}^2(\mathbf{x}, \mathbf{x}')$ with $c_0 < 1$, the previous equation yields:

$$\begin{aligned}
& |(\mathbf{m}_{t_0,t}^{\mathbf{x}}(\mathbf{x}') - \boldsymbol{\theta}_{t_0,t}(\mathbf{x}'))_i| \\
& \leq C \left(c_0 \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}') + (c_0^i + c_0) \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}') + (c_0^{i-\frac{1}{2}} + c_0) \mathbf{d}^{2i-1}(\mathbf{x}, \mathbf{x}') \right),
\end{aligned}$$

using again $\mathbf{d}(\mathbf{x}, \mathbf{x}') \leq 1$ for the middle term. After summing and by convexity inequalities, we eventually deduce:

$$\mathbf{d}(\mathbf{m}_{s,t}^{\mathbf{x}}(\mathbf{x}'), \boldsymbol{\theta}_{s,t}(\mathbf{x}')) \leq C c_0^{\frac{1}{2n-1}} \mathbf{d}(\mathbf{x}, \mathbf{x}').$$

This concludes the proof of Lemma 12. \square

B Parametrix expansion with different freezing points

In this section we show how the parametrix expansion (3.21) involving different freezing points can be derived. This can actually been done from the Duhamel formulation up to an additional discontinuity term. Restarting from (2.20) we can indeed rewrite from the Markov property that for given $(t, \mathbf{x}') \in [0, T] \times \mathbb{R}^{nd}$ and any $r \in (t, T], \boldsymbol{\xi}' \in \mathbb{R}^{nd}$:

$$u(t, \mathbf{x}') = \left[\tilde{P}_{r,t}^{\boldsymbol{\xi}'} u(r, \cdot) \right](\mathbf{x}') + \int_t^r ds \left[\tilde{P}_{s,t}^{\boldsymbol{\xi}'} f(s, \cdot) \right](\mathbf{x}') + \int_t^r ds \left[\tilde{P}_{s,t}^{\boldsymbol{\xi}'} \left((L_s - \tilde{L}_s^{\boldsymbol{\xi}'}) u \right) (s, \cdot) \right](\mathbf{x}'). \quad (\text{B.1})$$

Differentiating the above expression in $r \in (t, T]$ yields for any $\boldsymbol{\xi}' \in \mathbb{R}^{nd}$:

$$0 = \partial_r \left[\tilde{P}_{r,t}^{\boldsymbol{\xi}'} u(r, \cdot) \right](\mathbf{x}') + \left[\tilde{P}_{r,t}^{\boldsymbol{\xi}'} f(r, \cdot) \right](\mathbf{x}') + \left[\tilde{P}_{r,t}^{\boldsymbol{\xi}'} \left((L_r - \tilde{L}_r^{\boldsymbol{\xi}'}) u \right) (r, \cdot) \right](\mathbf{x}'). \quad (\text{B.2})$$

Denoting by $t_0 \in (t, T]$ the time at which we change the freezing point and integrating (B.2) on $[t, t_0]$ for a first given $\boldsymbol{\xi}'$ and between $[t_0, T]$ with a possibly different $\tilde{\boldsymbol{\xi}}'$ yields:

$$\begin{aligned}
0 & = \left[\tilde{P}_{t_0,t}^{\boldsymbol{\xi}'} u(t_0, \cdot) \right](\mathbf{x}') - u(t, \mathbf{x}') + \int_t^{t_0} ds \left[\tilde{P}_{s,t}^{\boldsymbol{\xi}'} f(s, \cdot) \right](\mathbf{x}') + \int_t^{t_0} ds \left[\tilde{P}_{s,t}^{\boldsymbol{\xi}'} \left((L_s - \tilde{L}_s^{\boldsymbol{\xi}'}) u \right) (s, \cdot) \right](\mathbf{x}') \\
& \quad + \left[\tilde{P}_{T,t_0}^{\tilde{\boldsymbol{\xi}}'} u(T, \cdot) \right](\mathbf{x}') - \left[\tilde{P}_{t_0,t_0}^{\tilde{\boldsymbol{\xi}}'} u(t_0, \cdot) \right](\mathbf{x}') + \int_{t_0}^T ds \left[\tilde{P}_{s,t_0}^{\tilde{\boldsymbol{\xi}}'} f(s, \cdot) \right](\mathbf{x}') + \int_{t_0}^T ds \left[\tilde{P}_{s,t_0}^{\tilde{\boldsymbol{\xi}}'} \left((L_s - \tilde{L}_s^{\tilde{\boldsymbol{\xi}}'}) u \right) (s, \cdot) \right](\mathbf{x}').
\end{aligned}$$

Recalling that $u(T, \cdot) = 0$ (terminal condition), the above equation rewrites:

$$\begin{aligned} u(t, \mathbf{x}') &= \int_t^T ds \left(\mathbb{I}_{s \leq t_0} \left[\tilde{P}_{s,t}^{\xi'} f(s, \cdot) \right] (\mathbf{x}') + \mathbb{I}_{s > t_0} \left[\tilde{P}_{s,t}^{\tilde{\xi}'} f(s, \cdot) \right] (\mathbf{x}') \right) \\ &\quad + \left[\tilde{P}_{t_0,t}^{\xi'} u(t_0, \cdot) \right] (\mathbf{x}') - \left[\tilde{P}_{t_0,t}^{\tilde{\xi}'} u(t_0, \cdot) \right] (\mathbf{x}') \\ &\quad + \int_t^T ds \left(\mathbb{I}_{s \leq t_0} \left[\tilde{P}_{s,t}^{\xi'} \left((L_s - \tilde{L}_s^{\xi'}) u \right) (s, \cdot) \right] (\mathbf{x}') + \mathbb{I}_{s > t_0} \left[\tilde{P}_{s,t}^{\tilde{\xi}'} \left((L_s - \tilde{L}_s^{\tilde{\xi}'}) u \right) (s, \cdot) \right] (\mathbf{x}') \right). \end{aligned}$$

We see that for $\xi' \neq \tilde{\xi}'$ we have an additional discontinuity term deriving from the change of freezing point along the time variable. Eventually, the above equation precisely gives (3.21), recalling that for $t_0 = t + c_0 |(\mathbf{x}_i - z)|^{2/(2i-1)}$, $\mathbb{I}_{s \leq t_0} = \mathbb{I}_{S_i}$.

C Appendix: Some reminders about Besov spaces

C.1 Thermic characteristic of the Besov space

Let us now recall some definitions/characterizations from Section 2.6.4 of Triebel [Tri83]. For $\alpha \in \mathbb{R}, q \in (0, +\infty], p \in (0, \infty]$, $B_{p,q}^\alpha(\mathbb{R}^d) := \{f \in \mathcal{S}'(\mathbb{R}^d) : \|f\|_{\mathcal{H}_{p,q}^\alpha} < +\infty\}$ where $\mathcal{S}(\mathbb{R}^d)$ stands for the Schwartz class and

$$\|f\|_{\mathcal{H}_{p,q}^\alpha} := \|\varphi(D)f\|_{L^p(\mathbb{R}^d)} + \left(\int_0^1 \frac{dv}{v} v^{(m-\frac{\alpha}{2})q} \|\partial_v^m h_v \star f\|_{L^p(\mathbb{R}^d)}^q \right)^{\frac{1}{q}}, \quad (\text{C.1})$$

with $\varphi \in C_0^\infty(\mathbb{R}^d)$ (smooth function with compact support) is s.t. $\varphi(0) \neq 0$, $\varphi(D)f := (\varphi \hat{f})^\vee$ where \hat{f} and $(\varphi \hat{f})^\vee$ respectively denote the Fourier transform of f and the inverse Fourier transform of $\varphi \hat{f}$. The parameter m is an integer s.t. $m > \alpha/2$ and for $v > 0, z \in \mathbb{R}^d$, $h_v(z) := \frac{1}{(2\pi v)^{d/2}} \exp\left(-\frac{|z|^2}{2v}\right)$ is the usual heat kernel of \mathbb{R}^d . We point out that the quantities in (C.1) are well defined for $q < \infty$. The modifications for $q = +\infty$ are obvious and can be written passing to the limit.

Observe that the quantity $\|f\|_{\mathcal{H}_{p,q}^\alpha}$, where the subscript \mathcal{H} stands to indicate the dependence on the heat-kernel, depends on the considered function φ and the chosen $m \in \mathbb{N}$. It also defines a quasi-norm on $B_{p,q}^\alpha(\mathbb{R}^d)$. The previous definition of $B_{p,q}^\alpha(\mathbb{R}^d)$ is known as the thermic characterization of Besov spaces and is particularly well adapted to our current framework. By abuse of notation we will write as soon as this quantity is finite $\|f\|_{\mathcal{H}_{p,q}^\alpha} =: \|f\|_{B_{p,q}^\alpha}$.

C.2 Equivalence of Besov norms

Proposition 14. *There is a constant $C > 1$ such that for all $f \in B_{\infty,\infty}^\alpha(\mathbb{R}^d)$, $\alpha \in (0, 1)$ we have:*

$$C^{-1} \|f\|_{B_{\infty,\infty}^\alpha} \leq \|\mathbf{D}f\|_{B_{\infty,\infty}^{\alpha-1}} \leq C \|f\|_{B_{\infty,\infty}^\alpha}.$$

We write also that $\|f\|_{B_{\infty,\infty}^\alpha} \asymp \|\mathbf{D}f\|_{B_{\infty,\infty}^{\alpha-1}}$.

Proof of Proposition 14. To establish this equivalence of norms we use the local means characterisation of the Besov space, see for instance 2.5.3 in [Tri83]. In other words, for functions $k_0, k^0 \in C^\infty(\mathbb{R}^d)$ (infinitely differentiable) such that $\text{supp}(k_0), \text{supp}(k^0) \subset \mathbf{B}(\mathbf{0}_{d,d}, 1)$ (the unit ball centered at the origin) and introducing for all $x \in \mathbb{R}^d$:

$$k(x) = \sum_{i=1}^d \partial_{x_i}^2 k^0(x),$$

we define the local means for all $(t, x) \in (0, +\infty) \times \mathbb{R}^d$ as:

$$k(t, f)(x) := \int_{\mathbb{R}^d} dy k(y) f(x + ty), \quad k_0(t, f)(x) := \int_{\mathbb{R}^d} dy k_0(y) f(x + ty).$$

It is then known that for all $f \in B_{p,q}^\alpha(\mathbb{R}^d)$:

$$\|f\|_{B_{\infty,\infty}^\alpha} \asymp \|k_0(1, f)(\cdot)\|_\infty + \sup_{t \in (0,1)} (t^{-\alpha} \|k(t, f)(\cdot)\|_\infty).$$

We can consider similarly the sequences of functions, $k^i = \partial_{x_i} k$, $k_0^i = \partial_{x_i} k_0$, $i \in \llbracket 1, d \rrbracket$, which satisfy the same assumptions as k and k_0 . Hence, we can write:

$$\begin{aligned} \|f\|_{B_{\infty, \infty}^{\alpha}} &\asymp \|k_0^i(1, f)(\cdot)\|_{\infty} + \sup_{t \in (0, 1)} (t^{-\alpha} \|k^i(t, f)(\cdot)\|_{\infty}) \\ &\asymp \|k_0(1, \partial_{x_i} f)(\cdot)\|_{\infty} + \sup_{t \in (0, 1)} (t^{1-\alpha} \|k(t, \partial_{x_i} f)(\cdot)\|_{\infty}) \\ &\asymp \|\partial_{x_i} f\|_{B_{\infty, \infty}^{\alpha-1}}. \end{aligned}$$

The penultimate identity is a consequence of an integration by parts over the convolution products, and the last equivalence is obtained again thanks to Theorem 2.5.3 in [Tri83]. Iterating over each component yields the result. □

Acknowledgments.

For the first Author, this work has been partially supported by the ANR project ANR-15-IDEX-02. For the third author, the article was prepared within the framework of a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program.

References

- [AH96] D. R. Adams and L. I. Hedberg. *Function spaces and Potential Theory*. Springer, 1996.
- [Bas97] R. F. Bass. *Diffusions and Elliptic Operators*. Springer, 1997.
- [BC03] R. Bass and Z. Q. Chen. Brownian motion with singular drift. *Ann. Probab.*, 31–2:791–817, 2003.
- [BCLP10] M. Bramanti, G. Cupini, E. Lanconelli, and E. Priola. Global L^p estimates for degenerate Ornstein-Uhlenbeck operators. *Math. Z.*, 266(4):789–816, 2010.
- [BCLP13] M. Bramanti, G. Cupini, E. Lanconelli, and E. Priola. Global L^p estimates for degenerate Ornstein-Uhlenbeck operators with variable coefficients. *Math. Nachr.*, 286(11-12):1087–1101, 2013.
- [BZ11] M. Bramanti and M. Zhu. L^p and Schauder estimates for nonvariational operators structured on Hörmander vector fields with drift. *Analysis of PDEs*, 2011.
- [CdR17] P.-E. Chaudru de Raynal. Strong existence and uniqueness for degenerate SDE with Hölder drift. *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, 53(1):259–286, February 2017.
- [CdR18] P.-E. Chaudru de Raynal. Weak regularization by stochastic drift: result and counter example. *Discrete and Continuous Dynamical Systems (Series A)*, 38–3:1269–1291, 2018.
- [CdRHM18] P.-E. Chaudru de Raynal, I. Honoré, and S. Menozzi. Sharp Schauder Estimates for some Degenerate Kolmogorov Equations. working paper or preprint, 2018.
- [CdRM17] P.-E. Chaudru de Raynal and S. Menozzi. Regularization effects of a noise propagating through a chain of differential equations: an almost sharp result. *Preprint*, <https://hal.archives-ouvertes.fr/hal-01613679>, 2017.
- [CG16] R. Catellier and M. Gubinelli. Averaging along irregular curves and regularisation of ODEs. *Stoch. Proc and Appl.*, 126–8:2323–2366, 2016.
- [DF14] F. Delarue and F. Flandoli. The transition point in the zero noise limit for a 1d Peano example. *Discrete and Continuous Dynamical Systems*, 34(10):4071–4083, April 2014.
- [DM10] F. Delarue and S. Menozzi. Density estimates for a random noise propagating through a chain of differential equations. *Journal of Functional Analysis*, 259–6:1577–1630, 2010.
- [DPL89] R. Di Perna and P. L. Lions. Ordinary Differential Equations, transport theory and Sobolev Spaces. *Inv. Math.*, 98:511–547, 1989.

- [FF11] E. Fedrizzi and F. Flandoli. Pathwise Uniqueness and Continuous Dependence for SDEs with Nonregular Drift. *Stochastics*, 83-3:241–257, 2011.
- [FFPV17] E. Fedrizzi, F. Flandoli, E. Priola, and J. Vovelle. Regularity of stochastic kinetic equations. *Electronic Journal of Probability*, 22:# 48, 2017.
- [Fla11] F. Flandoli. *Random perturbation of PDEs and fluid dynamic models*, volume 2015 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2011. Lectures from the 40th Probability Summer School held in Saint-Flour, 2010.
- [Fri64] A. Friedman. *Partial differential equations of parabolic type*. Prentice-Hall, 1964.
- [GO13] M. Gradinaru and Y. Offret. Existence and asymptotic behaviour of some time-inhomogeneous diffusions. *Annales Instit. H. Poincaré, Probabilités et Statistiques*, 49(1):182–207, 2013.
- [HM16] L. Huang and S. Menozzi. A Parametrix Approach for some Degenerate Stable Driven SDEs. *Annales Instit. H. Poincaré*, 52(4):1925–1975, 2016.
- [Hör67] L. Hörmander. Hypoelliptic second order differential equations. *Acta Mathematica*, 119:147–171, 1967.
- [Kol34] A. N. Kolmogorov. Zufällige Bewegungen (zur Theorie der Brownschen Bewegung). *Ann. of Math.*, 2-35:116–117, 1934.
- [KR05] N. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Prob. Theory Rel. Fields*, 131:154–196, 2005.
- [Lor05] L. Lorenzi. Estimates of the derivatives for a class of parabolic degenerate operators with unbounded coefficients in \mathbb{R}^N . *Annali della Scuola Normale Superiore di Pisa. Classe di Scienze. Serie V*, 2, 01 2005.
- [LR02] P.-G. Lemarié-Rieusset. *Recent developments in the Navier-Stokes problem*. CRC Press, 2002.
- [Men11] S. Menozzi. Parametrix techniques and martingale problems for some degenerate Kolmogorov equations. *Electronic Communications in Probability*, 17:234–250, 2011.
- [Men18] S. Menozzi. Martingale problems for some degenerate Kolmogorov equations. *Stoc. Proc. Appl.*, 128-3:756–802, 2018.
- [MS67] H. P. McKean and I. M. Singer. Curvature and the eigenvalues of the Laplacian. *J. Differential Geometry*, 1:43–69, 1967.
- [Pri15] E. Priola. On weak uniqueness for some degenerate SDEs by global L^p estimates. *Potential Anal.*, 42(1):247–281, 2015.
- [Str08] D. W. Stroock. *Partial differential equations for probabilists*, volume 112 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2008.
- [SV79] D.W. Stroock and S.R.S. Varadhan. *Multidimensional diffusion processes*. Springer-Verlag Berlin Heidelberg New-York, 1979.
- [Tri83] H. Triebel. *Theory of function spaces, II*. Birkhauser, 1983.
- [Ver80] A. Y. Veretennikov. Strong solutions and explicit formulas for solutions of stochastic integral equations. *Matematicheskii Sbornik. Novaya Seriya*, 111(153)(3):434–452, 480, 1980.
- [Ver83] A. Y. Veretennikov. Stochastic equations with diffusion that degenerates with respect to part of the variables. *Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya*, 47(1):189–196, 1983.
- [WZ16] F.Y. Wang and X. Zhang. Degenerate SDE with Holder-Dini drift and non-Lipschitz noise coefficient. *SIAM J. Math. Anal.*, 48(3):2189–2226, 2016.
- [Zha10] X. Zhang. Well-posedness and large deviation for degenerate SDEs with Sobolev coefficients. *Revista Matemática Iberoamericana*, 29(1), 2010.

- [Zha18] X. Zhang. Stochastic Hamiltonian flows with singular coefficients. *Sci. China Math.*, 61(8):1353–1384, 2018.
- [Zvo74] A. K. Zvonkin. A transformation of the phase space of a diffusion process that will remove the drift. *Mat. Sb. (N.S.)*, 93(135):129–149, 152, 1974.